PHOTOCHEMICAL SYSTEMATIC EVOLUTION OF LIGANDS BY EXPONENTIAL ENRICHMENT (PHOTOSELEX) EMPLOYING THE 5-BROMO-2'-DEOXYURIDINE CHROMOPHORE FOR THE BASIC FIBROBLAST GROWTH FACTOR TARGET

by

MACE C. GOLDEN

B.S., University of Washington, Seattle, WA 1987

M.S., Colorado State University, Ft. Collins, CO 1992

A thesis submitted to the

Faculty of the Graduate School of the

University of Colorado in partial fulfillment

of the requirement for the degree of

Doctor of Philosophy

Department of Chemistry and Biochemistry

1999

REPORT DOCUMENTATION PAGE			Form Approved OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	20.Aug.99	DISSERTATION

4. TITLE AND SUBTITLE
PHOTOCHEMICAL SYSTEMATIC EVOLUTION OF LIGANDS BY EXPONENTIAL ENRICHMENT (PHOTOSELEX) EMPLOYING THE 5-BROMO-2'-DEOXYURIDINE CHROMOPHORE FOR THE BASIC

5. FUNDING NUMBERS

6. AUTHOR(S)
CAPT GOLDEN MACE L

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
UNIVERSITY OF COLORADO AT BOULDER

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
THE DEPARTMENT OF THE AIR FORCE
AFIT/CIA, BLDG 125
2950 P STREET
WPAFB OH 45433

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
FY99-263

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION AVAILABILITY STATEMENT
Unlimited distribution
In Accordance With AFI 35-205/AFIT Sup 1

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

14. SUBJECT TERMS

15. NUMBER OF PAGES
136

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT

This thesis entitled:
Photochemical Systematic Evolution of Ligands by Exponential Enrichment (PhotoSELEX) Employing the 5-Bromo-2'-deoxyuridine Chromophore for the Basic Fibroblast Growth Factor Target
has been approved for the Department of Chemistry and Biochemistry

Tad H. Koch

Randy L. Halcomb

Date April 22, 1999

The final copy of this thesis has been examined by the signators, and we find that both the content and the form meet acceptable presentation standards of scholarly work in the above mentioned discipline.

Golden, Mace C. (Ph.D., Chemistry)

Photochemical Systematic Evolution of Ligands by Exponential Enrichment (PhotoSELEX) Employing the 5-Bromo-2'-deoxyuridine Chromophore for the Basic Fibroblast Growth Factor Target

Thesis directed by Professor Tad H. Koch

Abstract: An in vitro selection methodology (PhotoSELEX) was investigated and used to identify ssDNA aptamers capable of binding with high affinity and capable of covalently cross-linking via a high quantum yield photochemical reaction the human basic Fibroblast Growth Factor (bFGF) protein. Beginning with a library of ssDNA aptamers in which thymidine had been replaced with the chromophore 5-bromo-2'-deoxyuridine, five rounds of affinity-based in vitro selection were completed. The resulting library was then irradiated with 308 nm light in the presence of bFGF. Those sequences capable of forming a covalent bond to the target protein were subsequently partitioned from those not capable of forming such a bond. Enzymatic digestion of the protein component yielded a template suitable for polymerase chain reaction (PCR) amplification permitting additional rounds of such photocross-linking selection. Systematic and exclusive application of rounds of PhotoSELEX on such a library identified ssDNA aptamers with high affinity binding and high quantum yield cross-linking potential for the target. Moreover, it was shown that PhotoSELEX can be manipulated to yield 1) exceptionally high affinity binding or 2) exceptionally high quantum yield cross-linking aptamers depending on the nature of the pressure applied during the experiment. Based upon these results, a PhotoSELEX experiment was performed in which photochemical and affinity pressure were applied simultaneously to a completely random oligonucleotide library. This experiment showed that

PhotoSELEX can converge to unique sequences much more rapidly than affinity-based SELEX. Finally, the application of electrospray mass spectrometry (ESMS) and Maxam-Gilbert sequencing as novel methods to characterize the results of covalent-bond based SELEX protocols were demonstrated.

DEDICATION

Try as I might, I found unable to stand alone during the completion of this, the most arduous and all-consuming adventure in my life. While many have contributed mightily to my ultimate success, no one's contribution compares to the sacrifices made by my wife, Karla. She is, without question, the most intelligent and talented person I have ever known; indeed, several wonderful opportunities were offered to her during my time in this program. Yet, she stayed with me throughout and supported me in ways which are too numerous to mention. This degree is nearly as much hers as mine; I will forever be grateful for her unfailing devotion.

ACKNOWLEDGEMENTS

I have never engaged in any task for which I depended on so many for so much as I have for the completion of my doctoral degree. I am, first and foremost, indebted to my wife for continuous support and sacrifice. Beyond this, I would like to thank my advisor, Professor Tad Koch, for his motivation, compassion and ambition throughout the entire program. Mr. Brian Collins of NexStar Pharmaceuticals personally taught me nearly every biochemical laboratory technique I employed in successfully completing my doctoral research. Without his extreme patience and amazing good humor, this project would have been beyond my ability. I would also like to thank Ms. Hilary Opperman, the Graduate Secretary in the Department of Chemistry and Biochemistry. She employed her quick thinking and vast knowledge of arcane graduate school rules to bale me out of inextricable binds on more than one occasion. I also owe a debt of gratitude to the United States Air Force for the confidence they exhibited and support they provided me to complete this program. In particular, the Air Force Institute of Technology paid my tuition and provided other financial assistance. Also, the Department of Chemistry at the United States Air Force Academy dedicated every conceivable resource imaginable to help guarantee my success. Namely, Col Hans Mueh, Col Clifford Utermoehlen, LtC Ron Furstenau, Dr. Don Bird, Dr. Norman Heimer, Dr. Joel Cain, Dr. Barry Hicks, Maj Matt Morgan, Maj Luke Lorang, Mrs. Liz Curry, and Mr. Ike Sleighter made invaluable contributions to the disseratation and its defense. I am especially indebted to Mr. Sleighter for his tireless assistance and unequalled ability with Microsoft Power Point. Finally, I would like to thank my family. My father and mother, Leon

and Verjean Golden, taught me the value of persistence. My in-laws, Jan and Karl Clark, taught me that family comes first. Everyone, especially Alyxandra, Tristan and Tessa, taught me how good it feels to be loved.

CONTENTS

LIST OF FIGURES

Figure

LIST OF TABLES

Table

LIST OF SCHEMES

Scheme

LIST OF ABBREVIATIONS

PhotoSELEX	Photochemical Systematic Evolution of Ligands by Exponential Enrichment
ss DNA	single-stranded deoxyribonucleic acid
bFGF	basic Fibroblast Growth Factor
SELEX	Systematic Evolution of Ligands by Exponential Enrichment
ESMS	electrospray mass spectrometry
PCR	polymerase chain reaction
VEGF	Vascular Endothelial Growth Factor
SPS	sodium dodecyl sulfate
AMV	avian myeloblastosis virus
K_d	dissociation constant
BrdU	5-bromo-2'-deoxyuridine
BrU	5-bromouracil
IU	5-iodouracil
ECM	extracellular matrix
HSPG	heparan sulfate proteoglycan
PG	proteoglycan
PBS	phosphate buffered saline
#xA	number of rounds of affinity SELEX through which a library has evolved
#xP	number of rounds of PhotoSELEX through which a library has evolved
XLBB	cross-linking binding buffer

XL	covalently-bound oligonucleotide/protein adduct (cross-link)
XL-dig	cross-linked adduct with bFGF component digested with trypsin
DTT	dithiothreitol
SVP	snake venom phosphodiesterase
AP	alkaline phosphatase

CHAPTER I: INTRODUCTION

SELEX: An affinity association-based combinatorial chemistry technique

Combinatorial chemistry has emerged as an important explorative technique in the pharmaceutical industry. Rather than investigating tens of compounds annually as is permitted using traditional chemical approaches such as random screening or rational drug design, combinatorial chemistry technology facilitates evaluation of enormous numbers of potential compounds for catalytic, diagnostic or therapeutic properties (Gold, 1995). While several different combinatorial methodologies exist, they all share the fundamental principle of screening a library of molecules for specific properties and then selecting those with the most desirable attributes. Such combinatorial chemistry methodologies include protein/peptide-based libraries among others. Probably the best combinatorial chemistry technology currently employed by man is the human immune system. The strength of the immune system lies in its ability to generate a huge diversity of solutions to a given target ***and*** its ability to access/select the very best performers for the desired task (usually high-affinity binding). Another combinatorial chemistry technique currently in use is the Systematic Evolution of Ligands by Exponential Enrichment (SELEX) (Tuerk and Gold, 1990; Gold et al., 1995).

In SELEX, the large diversity of solutions is obtained by generating a library of random-sequence oligonucleotides (RNA, ssDNA or oligonucleotides with modified bases). Long held dogma views nucleic acids as carriers of information and proteins as carriers of conformation--tapes vs. shapes (Judson, 1979). However, elucidation of the structure of tRNA and the discovery of RNA's catalytic activity in 1982 proved that nucleic acids indeed possess interesting chemical properties

resulting from their three-dimensionality (Kim et al., 1974; Gold et al., 1995). Thus, a library of single-stranded oligonucleotides should possess a wealth of conformational and chemical diversity. The initial library can be generated by solid-phase oligonucleotide synthesis. A prototype member of the library, shown below, has fixed regions at both ends of the molecule which flank a random region. As will be discussed later, the fixed regions permit primer annealing and subsequent

5'-xxxxxxxxxxxxxxxNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNxxxxxxxxxxxxxxx-3'

amplification by the polymerase chain reaction (PCR). During the synthesis of the library's random region, an equimolar concentration of the four base phosphoramidites is introduced during each linkage of the growing DNA chain. Each base then has a nearly equal probability of linking to any given chain; consequently, a library of many different oligonucleotide molecules is produced. The diversity of the resulting library can be increased by increasing the length of the random region since the total number of different molecules that can be generated is 4^N where N is the number of base positions. For a 30-mer random region, this gives 1.153×10^{18} different possible molecules. Gold (1995) argues that most of the known simple single-stranded oligonucleotide motifs including hairpins, bulges within helices, psuedoknots and G-quartets are attainable with random regions as small as 30 bases. Moreover, as drug identification is frequently the goal of a SELEX experiment, short olgonucleotides are preferred so random regions significantly longer than 30 bases are seldom employed.

Traditional, and still the work-horse protocol of SELEX technology, relies on identifying oligonucleotide molecules which bind a target molecule with high-affinity and specificity. Initially, researchers employing the SELEX technology selected

known nucleic acid binding proteins as SELEX targets. While such proteins have been successfully targeted and high-affinity binding ligands identified, subsequent SELEX experiments have found that non-nucleic acid binding protein targets such as growth factors, proteases, antibodies and small peptides are suitable substrates as well (Gold, 1995; Gold et al., 1995). With these successes, researchers experimented with and found high-affinity binding ligands for a wide array of small molecule targets including redox sensitive molecules, dyes, drugs, amino acids and other nucleotides (Gold, 1995; Ellington and Szostak, 1990; Osborne and Ellington, 1997). Surprisingly, few extremely high-affinity SELEX-derived oligonucleotides have been reported for nucleic acid binding proteins, but several such aptamers have been reported for non-nucleic acid binding proteins (Ruckman et al., 1998; Green et al., 1996; Schneider et al., 1995; Ellington and Szostak, 1990).

In addition to successfully identifying scores of ligands which bind with high-affinity their protein targets, SELEX has succeeded in identifying ligands which exhibit extraordinary specificity. For example, SELEX-derived ligands for Vascular Endothelial Growth Factor (VEGF) bind its heparin-binding site but do not bind with other members of the growth factor family even though they all bind heparin (Jellinek et al., 1993). In other experiments SELEX has evolved aptamers which readily distinguish theophylline from caffeine (Jenison et al., 1994) and certain amino acids from their enantiomers (Famoluk and Szostak, 1992).

Gold (1995) argues the success of SELEX-derived oligonucleotides may result from their constrained, unique three-dimensional structures. This, he contends, permits near perfect molecular partners to take advantage of complementary shapes, ionic interactions (as well as other intermolecular forces to include London forces),

and hydrophobicity. Gold also states that perfect binding partners are a rare occurence in biological systems as finite associations are required for organismic evolution. In essence, then, SELEX is a manipulation of biological molecules which would have self-eliminated evolutionarily. Although the oligonucleotides identified through SELEX exhibit impressive sensitivity and selectivity for their targets, the affinity-based SELEX methodology is surprisingly simple in practice. To begin, the randomized library is incubated with the target molecule (assumed to be a protein for this discussion) in a suitable medium (e.g. phosphate buffered saline). This initial exposure allows those few sequences with special conformations, hydrogen bonding potential and hydrophobicity to specifically bind the target molecule. Next, those unique binding sequences in the library are partitioned from the vast majority which do not bind.

Partitioning is a crucial step and its efficiency in SELEX determines the rate of convergence. Partitioning is frequently accomplished via nitrocellulose filter partitioning although other partitioning techniques such as immobilization on a solid matrix or immunoprecipitation of the target are also employed (Fitzwater and Polisky, 1996). Nitrocellulose filter paper binds protein molecules nonspecifically and tenaciously while it binds oligonucleotides poorly if at all. As such, when the reaction medium is passed through the filter, only those oligonucleotides which have formed a relatively long-lived association with a protein molecule remain behind; nearly all others pass through. The remaining protein/oligonucleotide pairs are then eluted from the filter and each other with a strong protein-denaturing agent such as urea or formamide and sodium dodecyl sulfate. Finally, the partitioned oligonucleotides are amplified and prepared for the next round of selection.

The winning oligonucleotides are amplified via the polymerase chain reaction (PCR). Using primer oligonucleotides (short oligonucleotide sequences which can anneal via Watson-Crick base pairing to the fixed regions incorporated into each sequence of the library), each recovered sequence will be copied and amplified with a thermally-stable DNA polymerase. (When employing RNA libraries, the partitioned sequences are first converted to cDNA via the avian myeloblastosis virus (AMV) or other reverse transcriptase which subsequently can be amplified by PCR. Finally, before the next round of selection, the DNA amplified library is transcribed back to RNA using T7 RNA polymerase.) This process constitutes one round of SELEX; additional rounds are performed until convergence is achieved.

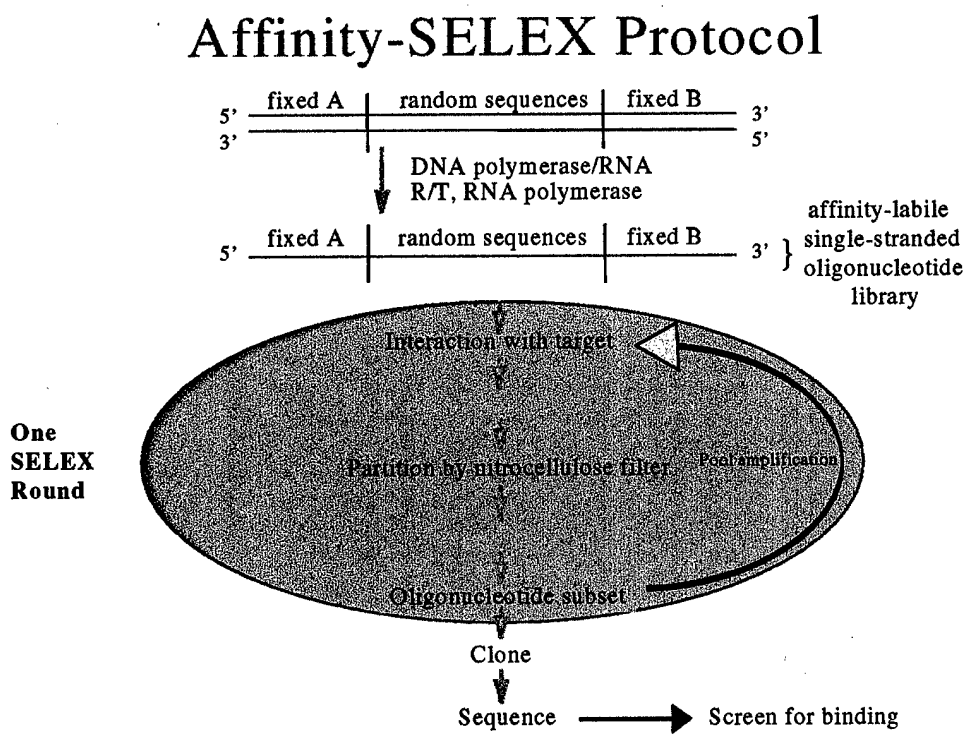


Figure I.1.1. The affinity-based SELEX protocol.

In affinity-based SELEX, the objective is to identify high-affinity binding aptamers for the protein target. Therefore, the progress of the SELEX can be monitored by measuring the dissociation constant (K_d) of each library. When no improvement in K_d is observed between one library and the next, the SELEX is considered completely evolved. The final library can then be cloned and sequenced to identify its individual members. These individual sequences can then be screened to determine those with the best binding for the protein target. A schematic for the affinity-association based SELEX protocol is shown in Figure I.1.1 on the previous page.

PhotoSELEX: A covalent bond-based combinatorial chemistry technique

Clearly, affinity-based SELEX has generated high-affinity binding aptamers to a wide array of molecules including many proteins. Moreover, as experience with oligonucleotide combinatorial libraries and their chemical properties grows, so, presumably, will the ability of researchers to identify higher and higher-affinity aptamers for a given target. The ultimate in affinity binding would be an aptamer/target association which never dissociated. Of course, such an association can be effected by the formation of a covalent bond between the aptamer and the target; the implication for diagnostics and therapeutics for such covalent adducts is immense.

One way to form a covalent bond between an aptamer and a target protein is with the use of modifed bases capable of acting as photoreactive chromophores. As will be discussed in the following section, halogenated uracil bases (among many other modified bases) can absorb UV light with subsequent electronic excitation and covalent bond formation to aromatic or sulfur-bearing amino acid residues in suitable

proximity and conformational orientation. The PhotoSELEX methodology would then be similar to the previously described affinity SELEX methodology except that selection would be based on the ability to form a covalent bond rather than mere high-affinity association.

PhotoSELEX would begin with preparation of the halogenated uridine-substituted library followed by combination with the target protein. This mixture would then be irradiated with appropriate wavelength light to form a cross-link to the protein with the few sequences from the library capable of forming such a bond. One of the fundamental hypotheses of this research is that the conformational orientation required between an aptamer and its protein target to permit photoinduced bond formation is much more demanding than that required for affinity association. As such, it is conceivable that PhotoSELEX will converge to a completely evolved library much more rapidly than affinity SELEX. However, as with affinity SELEX, the rate of convergence depends strongly on the ability to partition the desired sequences from the undesired ones.

While partitioning in affinity SELEX is accomplished with nitrocellulose filter binding, this is completely insufficient in PhotoSELEX as this technique would not only partition cross-linking sequences but affinity-associated sequences as well. Consequently, the PhotoSELEX partioning step is accomplished by denaturing urea or sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). Since the rate of migration of molecules in PAGE is primarily a function of their molecular weight, a covalent oligonucleotide-protein adduct will migrate more slowly than noncross-linked oligonucleotide sequences (affinity-based associations will be destroyed). Another fundamental hypothesis of this research is that PAGE

partitioning is significantly more effective than nitrocellulose partitioning. Again, this should substantially increase the rate of PhotoSELEX convergence vis a vis affinity SELEX. Once the cross-linking aptamers are partitioned from the noncross-linking aptamers, they must be recovered for amplification.

In affinity SELEX, recovery of the winning aptamers from the nitrocellulose filter is accomplished by simply denaturing the oligonucleotide sequence from the protein. This is again insufficient in PhotoSELEX. To recover the oligonucleotide sequence in this case, the protein is enzymatically digested with proteinase K or other suitable enzyme after elution from the gel. This step is crucial for the success of PhotoSELEX as insufficient digestion will leave the oligonucleotide so perturbed as to be unusable as a DNA polymerase template. Once recovered, the winning oligonucleotide sequences from the previous round can be amplified to permit the next round of PhotoSELEX. Additional rounds of PhotoSELEX can be performed until no further improvement from one library to the next is observed. In the case of PhotoSELEX, evolution can be monitored by measuring the library binding affinity and by measuring the library cross-linking yields relative to DNA through phosphor-imaged scans of PAGE. A PhotoSELEX experiment will be deemed complete when no increase in library binding affinity or cross-linking yield is observed. A schematic showing the putative PhotoSELEX protocol appears in Figure I.2.1 on the following page.

Jensen and coworkers (1995) performed the first experiment (and to date the only published experiment) to identify oligonucleotides capable of covalently binding their target. They successfully completed a modified SELEX protocol using an RNA oligonucleotide library incorporating the photoreactive chromophore 5-IU for the

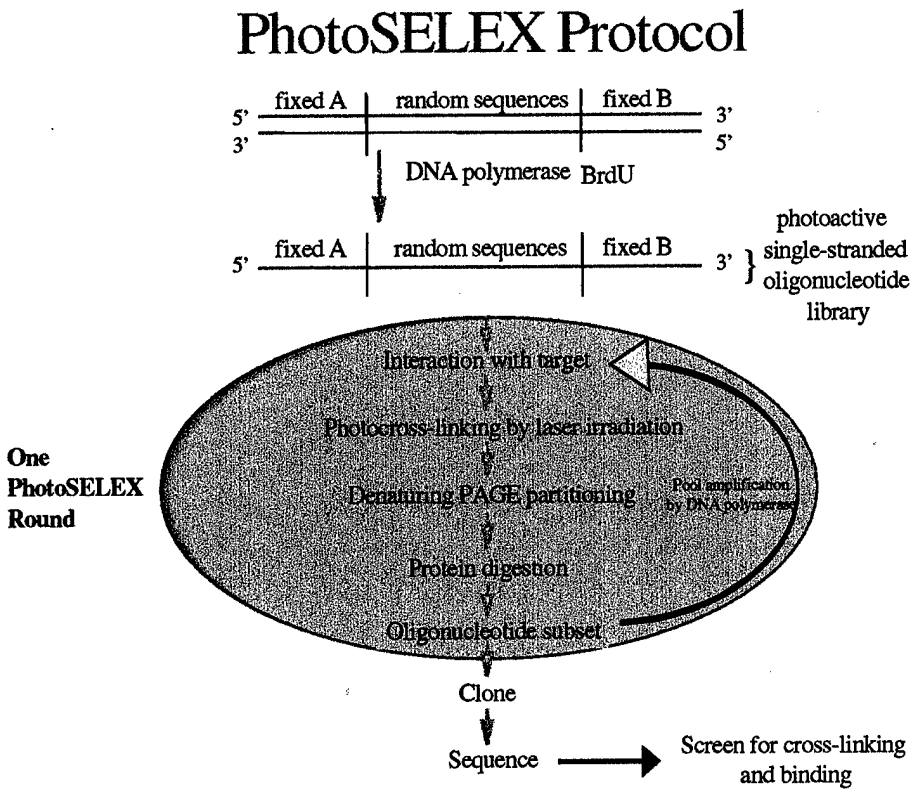


Figure I.2.1. The PhotoSELEX protocol.

human immunodefiency virus type 1 Rev protein target. HIV-1 Rev is a nucleic acid binding protein, and its high-affinity epitope (the IIB stem) within the HIV-1 RNA genome (Rev Responsive Element-RRE) to which it binds has been identified (Heaphy et al., 1990; Karn et al., 1991; Cook et al., 1991). Moreover, affinity SELEX protocols have been completed against HIV-1 Rev, and the results appear similar to the IIB stem in primary sequence and secondary structure (Tuerk et al., 1993; Tuerk and McDougal-Waugh, 1993; Jensen et al., 1994). Based upon these results, Jensen and coworkers prepared a biased randomized template for the oligonucleotide library (Jensen et al., 1995).

These researchers prepared a biased randomized template based upon an

aptamer identified by Jensen et al., (1994)--the prototype aptamer. This was done by preparing DNA oligonucleotides of the form:

5'-CCCGGATCCTCTTTACCTCTGTGTGagatacagagtccacaaacgtgttctcaatgacccGGTCGGAAGGCCATCAATAGTCCC-3'

in which the lower case letters indicate the biased randomized region. In this region, a mixture of phosphoramidites was used during the synthesis in which 62.5% was the phosphoramidite corresponding to the nucleotide in the prototype aptamer. The remaining 37.5% was equally divided among the other three phosphoramidites (Jensen et al., 1995). It should be noted that this is a departure from the normal SELEX protocol as it is envisioned that the successful SELEX project (affinity or covalent) will address a novel target for which oligonucleotide binding sequence information is unknown.

Once Jensen and coworkers had assembled the 5-IU substituted RNA library, they initiated the SELEX experiment. First, they verified that the prototype sequence and its fully substituted 5-IU analog did not cross-link HIV-1 Rev. The fact that they observed no cross-linked adduct in either case and observed a two order of magnitude reduction in binding for the 5-IU analog immediately called into question the need for biased randomization of the starting oligonucleotide library. Next, they performed three initial rounds of affinity selection (rounds 1-3) using standard nitrocellulose filter partitioning. These rounds employed 20-100 nM RNA and 1-6 nM HIV-1 Rev in suitable binding buffer. By keeping RNA in excess of protein, they ensured binding competition and selection of high-affinity aptamers for subsequent photocross-linking rounds (Jensen et al., 1995).

After the first three rounds of affinity selection, Jensen and coworkers completed four rounds (rounds 4-7 overall) in which they selected ligands capable of

forming a photocross-link to HIV-1 Rev. In round 4, they used 308 nm light from a XeCl excimer laser to irradiate the RNA/protein mixture for 3 min. In round 5, they used 325 nm light provided by a HeCd laser to irradiate the mixture for 30 min. Then in rounds 6 and 7, they irradiated the mixture for 10 min and 1 min, respectively, with 325 nm light. To maximize the opportunity for cross-linking, Jensen and coworkers performed the photocross-linking rounds in protein excess (50-100 nM RNA to 200 nM HIV-1 Rev in rounds 4-6 and 500 nM HIV-1 Rev in round 7). In each of these rounds they used denaturing urea-PAGE to partition the protein cross-linking sequences from the noncross-linking sequences. After partitioning, they recovered the cross-linking oligonucleotide sequences by excising the RNA/protein adduct from the gel, eluting the adduct from the gel slice, and then digesting the protein from the adduct with proteinase K (Jensen et al., 1995). The recovered RNA aptamers were then reverse transcribed to cDNA, amplified by PCR, and finally transcribed to RNA. This process of cross-linking, partitioning, recovery, and amplification was then repeated in every round of photocross-linking selection.

After four rounds of photocross-linking, Jensen and coworkers completed three more rounds of affinity selection (rounds 8-10 overall) and 3 more rounds of photocross-linking selection (rounds 11-13 overall). They argue that affinity and photoselection were accomplished simultaneously in rounds 11-13 as a result of the introduction of 10 μM yeast tRNA as a competitive binder (K_d of yeast tRNA for HIV-1 Rev not reported) (Jensen et al., 1995).

After 13 rounds of alternating affinity and photoselection, Jensen and coworkers cloned and sequenced their evolved library. They found two highly conserved sequence families which appear in different reading frames within the

biased randomized region. Interestingly, these photocross-linking sequences differed markedly from the original prototype sequence. Again, this seems to indicate that biasing the original oligonucleotide library is not only unnecessary in PhotoSELEX but may mitigate its ability to find the highest-affinity binders and highest quantum yield photocross-linkers. Jensen and coworkers determined from computer folding programs that the first of these families shared a stem-loop structure while the second formed a stem with the 5' fixed region. Upon sequence screening, they observed family I sequences to bind HIV-1 Rev with K_ds ranging from 1-10 nM and to cross-link with yields of 30-40%. Conversely, they found family II sequences to bind HIV-1 Rev with K_ds ranging from 30-50 nM and to cross-link with yields of 60-70%. Finally, they tested a truncated member of family I (trunc 3) and found it to bind HIV-1 Rev with a $K_d = 0.8$ nM and to cross-link with a yield of 40%. Unfortunately, they also observed that trunc 3 cross-linked with HIV-2 Rev which indicates limited specificity (Jensen et al., 1995).

In the final analysis, Jensen et al. were able to identify RNA aptamers capable of binding with reasonably high-affinity and cross-linking with good yield a protein target. However, while their experiment essentially proved the concept of PhotoSELEX, it answered few of the questions regarding the true potential for PhotoSELEX to identify high-affinity binding/high-quantum yield cross-linking aptamers. For instance, they completed the photocross-linking rounds in protein excess. This not only failed to take advantage of competitive binding (it is very questionable whether tRNA competes for HIV-1 Rev binding-even at high concentration), but it failed to select the best photocross-linkers from average photocross-linkers. They also arbitrarily changed cross-linking conditions during the

SELEX to facilitate its evolution. Finally, as mentioned several times above, they completed the SELEX using a biased randomized oligonucleotide library. In view of these shortcomings, the present research set out to address the following questions regarding PhotoSELEX:

1. Can PhotoSELEX be accomplished against a target for which any advantage derived from a biased-randomized library is unavailable?

2. Can PhotoSELEX not only identify photocross-linking aptamers but select the highest quantum yield photocross-linking aptamers among all those which cross-link?

3. Can PhotoSELEX truly select for affinity while selecting for photocross-linking?

4. Can PhotoSELEX be performed in a systematic, reproducible manner without *in situ* modifications to the protocol?

If the answer to these questions is no, then PhotoSELEX will be relegated to the role of an intellectual curiosity. If, however, the answer to these questions is yes, PhotoSELEX will at a minimum augment and may even supplant affinity SELEX as a combinatorial chemistry technique for the high throughput identification of diagnostic and therapeutic agents. While the exact protocol devised to achieve PhotoSELEX convergence for the bFGF target may not be that ultimately used for other targets, certain fundamental principles should be generally applicable. With an affirmative answer to these questions, this research will also seek to develop new techniques for characterizing the photocross-linked adduct which may prove valuable to the pharmaceutical industry.

Photochemistry of the 5-bromo-2'-deoxyuridine chromophore

Researchers have irradiated nucleoprotein complexes for the purpose of forming covalent bond(s) at the putative points of contact for over thirty years. Wild

type oligodeoxynucleotides are thought to form a covalent bond between a base and an amino acid of the protein primarily via a one-photon excitation, but, under appropriate conditions, via a two-photon excitation. Examples of single-photon initiated photocross-linking include fd phage DNA to fd gene 5 protein (Paridiso et al., 1979), DNA to *E. coli* RNA polymerase (Harrison et al., 1982), DNA to DNA polymerase (Markowits, 1972), and *E. coli* DNA to bovine serum albumin (Smith 1964). Most yields reported for such systems range between 5 and 20%. Surprisingly, Merrill and co-workers (1988) and Williams and Konigsberg (1991) report cross-linking yields as high as 85%.

Conversely, oligonucleotides which incorporate modified bases capable of absorbing UV light may form the covalent bond exclusively via a one-photon excitation. The literature is replete with such modified base chromophores. These include 5-azidouracil, 8-azidoadenine, 8-azidoguanine, 5-iodocytosine, and 4-thiouracil among several others (Meisenheimer and Koch 1997). However, of all modified bases available, none more closely met the properties required for success in the PhotoSELEX project than 5-bromouracil. In addition to well characterized and robust photochemical properties to be discussed below, 5-bromo-2'-deoxyuridine is a good substitute for the native deoxynucleotide thymidine. One of the key properties required of the modified base by PhotoSELEX is ready incorporation of its deoxynucleotide triphosphate form into growing oligonucleotide strands by polymerase enzymes during PCR. As bromine has very nearly the same van der Waals radius as the methyl group at the 5-position in thymidine (1.95 Å to 2.00 Å, respectively) and nominally greater electronegativity (3.0 to 2.6, respectively), its compatability with polymerase enzymes is not surprising (Khalili et al., 1988). In

addition to its size and electronegativity, bromine confers useful photochemical properties to the uracil base.

Many nucleoprotein complex structures have been elucidated through photocross-linking of the 5-bromouracil chromophore shown in Figure I.3.1 below.

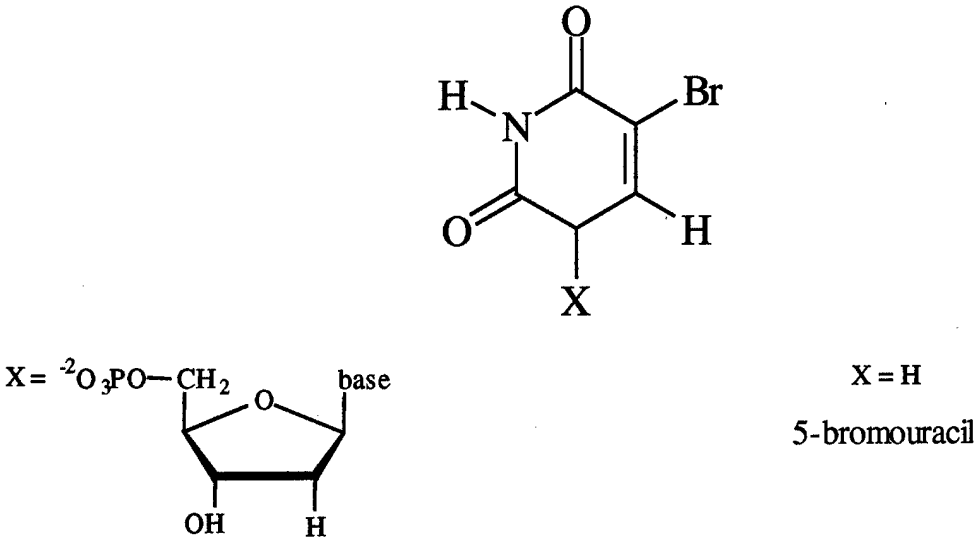


Figure I.3.1. The 5-bromouracil chromophore.

A few representative examples include lac repressor protein cross-linked to operator DNA (Allen et al., 1991), major chromosomal protein (MC1) and its DNA (Katousian-Safadi et al., 1991), and the R-17 bacteriophage coat protein and its associated viral RNA hairpin (Gott et al., 1991). Tanner et al. (1988) reported an impressive 50% cross-link yield for 5-bromouracil substituted presursor tRNA and yeast tRNA ligase. But to date, Katouzian-Safadi and Charlier (1994) have reported the highest cross-link yield for a 5-bromouracil-substituted nucleic acid/protein system when they observed 70% yield.

Figure I.3.2 below illustrates that the maximum absorbance of 5-bromouridine occurs at 280 nm. However, the absorbance band extends slightly beyond 300 nm; thus, 5-bromouridine can be excited by >300 nm light (Hicke et al., 1994).

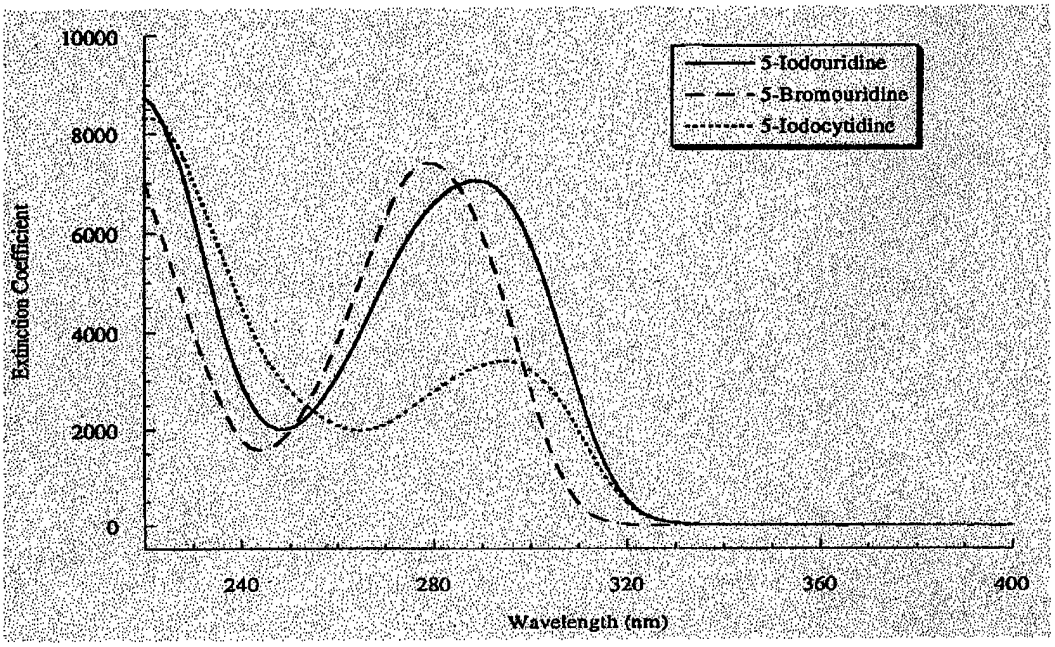


Figure I.3.2. UV absorbance spectrum of halogenated uracil bases (after Meisenheimer and Koch, 1997)

Moreover, Swanson et al. (1981) have observed the UV absorption spectrum of BrU taken in 2-propanol solvent to have a weak shoulder at 300 nm (ε 388 $M^{-1}cm^{-1}$) corresponding to an n-π^* transition. This relatively long-wavelength excitation characteristic of BrU is advantageous as it mitigates photolytic damage that shorter wavelength light may do to the rest of the nucleic acid, the protein or the chromophore itself.

Several research groups have extensively studied the photocross-linking mechanism of 5-bromouracil by evaluating the photoreduction of the chromophore

and by characterizing its cross-linking in model studies. In particular Ito et al. (1980) showed that photocoupling between 5-bromo-1,3-dimethyluracil and N^b-acetyltryptophan methyl ester was facilitated by the triplet sensitizer acetone. Conversely, these researchers found that sensitization of this photoreaction with acetophenone (E_T = 74 kcal/mol) and benzophenone (E_T = 69 kcal/mol) did not yield the coupled adduct even though these sensitizers presumably posses sufficient triplet energy to sensitize indoles, e.g. (E_T = 68 kcal/mol). From this they argue that the triplet state of BrU which is formed by energy transfer from triplet acetone (E_T = 79-82 kcal/mol) and not triplet indole is responsible for the photocoupling in this model system. Rothman and Kearns (1967) have estimated the energy of BrU's π,π^* T_1 as 74 kcal/mol above the ground state.

In addition, Ito and coworkers showed that irradiation of the BrU/tryptophan model system at 254 nm in the absence of sensitizer resulted primarily in photoreduction to yield uracil and only a small fraction of the photocoupled product. They also observed that adding 1,3-pentadiene under these circumstances served only to inhibit photocoupled product formation. From this they concluded that irradiation of BrU at 254 nm results in a short-lived excited singlet state which predominately undergoes C-Br bond homolysis in inefficient competition with inter-system crossing to the excited triplet. This provides additional support for the hypothesis that it is triplet BrU which forms during the photocoupling reaction.

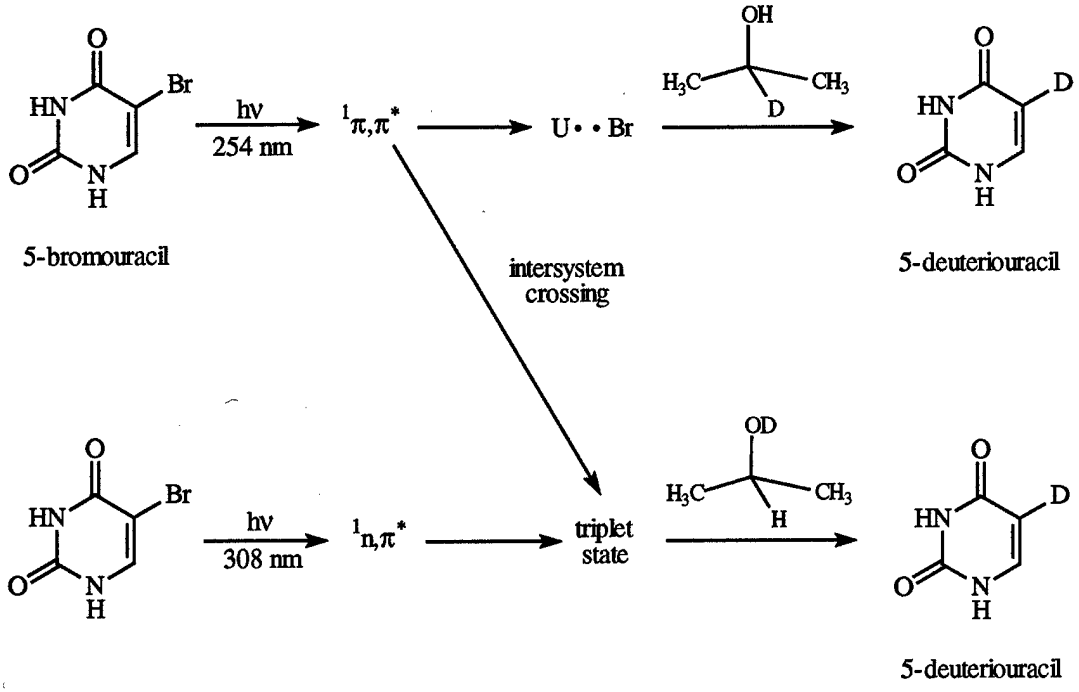
Finally, Ito and coworkers observed stringent regioselectivity in photocoupling to the 2-position of the indole ring. Indoles are known to undergo electrophilic substitution at the 3-position (Sundberg, 1970). Conversely, indoles engaged in radical chemistry are known to give a mixtrure of 1, 2, 3, 4, and 6-

substituted products (Somei and Natsume, 1973). However, electron transfer reactions such as photoaddition (Matsuo et al., 1976), anodic cyanation (Yoshida, 1977) and photo-sensitized electron transfer cyanation (Yoshida, 1979) all yield 2-substituted indoles as the primary product--a result consistent with charge distribution calculations which show the greatest positive charge at the 2 position. As a result, Ito and co-workers surmized that the photocoupling reaction between BrU and indole derivatives must proceed through an electron transfer mechanism.

In summary, these researchers hypothesised that the photocoupling of BrU to tryptophan derivatives must proceed through the BrU first excited triplet which associates with the indole to generate the triplet exciplex followed by electron transfer from tryptophan to BrU to yield the radical ion pair. $BrU^{-}\cdot$ subsequently dissociates into Br^{-} and the 5-uracilyl radical which may then covalently bond with the indole cation radical. Finally, this adduct is deprotonated to yield the photocoupled product.

The essential photochemical characterization of BrU as ascertained by Ito et al. has been largely corroborated by others analyzing the chromophore's photoreduction. Namely, Swanson et al. (1981) have examined the frequency-dependent results of deuterium incorporation from deuterium-labeled 2-propanol solvent and have confirmed Ito and coworkers essential conclusions. Additionally, Dietz et al. (1987) excited BrU at 254 nm in 2-deuterio-2-propanol solvent and observed uracil with 70% deuterium incorporation at the 5-position. Similar excitation in 2-propanol-d solvent yielded 30% deuterium incorporation at the 5-position of BrU. These researchers also completed deuterium incorporation studies at 308 nm light in 2-propanol-d solvent and observed 80% deuterated uracil. Dietz and coworkers have interpreted these results as indicating three excited states including

two excited singlets ($S_1 \rightarrow {}^1n,\pi^*$ and $S_2 \rightarrow {}^1\pi,\pi^*$) and one triplet ($T_1 \rightarrow {}^3\pi,\pi^*$). Evidently, the S_2 excited state may undergo C-Br bond homolysis resulting in a uracilyl radical which abstracts a hydrogen from the 2-position of 2-propanol. S_2 may also intersystem cross to T_1 and react with 2-propanol via an electron transfer mechanism. Conversely, S_1 may only intersystem cross to T_1. These conclusions are summarized in Scheme I.3.1 below.



Scheme I.3.1. UV promoted BrU deuterium incorporation from deuterated propanol.

While these results seemed to convincingly portray the photochemical mechanism of 5-bromouridine/tryptophan cross-linking in acetone and its photochemical reduction in deuterated propanol by identical excited states, several questions remained. For instance, Ito and coworkers found that acetone-sensitized

irradiation of isopropylidene-5-bromouridine and N-acetyltyrosine methyl ester did not produce the corresponding adduct. Also, no photocoupling between BrU and tryptophan, tyrosine or histidine was observed in aqueous solution (Saito et al., 1981; Saito and Matsuura, 1985). As a result, Dietz and Koch (1987) sought to further elucidate the mechanism of BrU photocross-linking by examining potential coupling with tryptophan, histidine, and tyrosine derivatives upon 308 nm light irradiation in aqueous solvent. Previously, these researchers achieved selective excitation of the $^1n,\pi^*$ first excited singlet state at this wavelength with subsequent intersystem crossing from S_1 to T_1 (Dietz et al., 1987). As such, Dietz and Koch argued that electron-donating aromatic amino acids should in general be capable of photocross-linking with BrU in an aqueous environment upon irradiation with 308 nm light.

These researchers did indeed observe photocoupling between BrU and all three amino acids. Although the observed quantum yields for the reactions were low (7.5×10^{-3}, 1.5×10^{-3}, 1.0×10^{-3}) for tryptophan, tyrosine and histidine respectively, which they attributed to a nonspecific association between the chromophore and amino acid derivative, they were consistent with known oxidative one-electron transfer rates and ionization potentials for the amino acids (Jovanovic et al., 1986; Amouoyal et al., 1979; Franklin et al., 1969; Ramsey, 1979). As a result, Dietz and Koch again concluded that the electron transfer mechanism originally proposed by Ito and coworkers was essentially correct and applied to all aromatic amino acid cross-linking reactions. Contrary to Ito and coworkers, Dietz and Koch do argue that bond formation occurs prior to release of Br^- from triplet BrU which minimizes the opportunity for back electron transfer.

While it is clear that these various model studies agree as to the

photochemical character of BrU, an examination of its behavior in a macromolecular system seems incongruous. Willis et al., (1994), studied R17 bacteriophage coat protein bound to various RNA stem loop structures incorporating BrU within the phage genome. They observed cross-linking between the 5-position of the RNA and Tyr85 of the coat protein. Unexpectedly, the maximum yield for the photocross-linked adduct (ca 40% based on RNA) was limited by photodegredation of the protein (hypothesized to result specifically from damage to tryptophan residues). Consequently, Norris et al. (1997) initiated an examination of BrU in RNA in an attempt to identify a reactive wavelength which would allow selective excitation of $^1n,\pi^*$ while simultaneously remaining transparent to tryptophan.

In their investigation, Norris and coworkers prepared the tyrosine derivative N-acetyltyrosine N-ethylamide which they irradiated in the presence of BrU in aqueous solvent. They noted strong frequency dependence of quantum yields for formation of cross-link and destruction of BrU when these quantum yields were based on excitation of BrU. Unexpectedly, the frequency dependence was eliminated when the quantum yields were based upon excitation of tyrosine. Norris and coworkers also observed that introduction of acetonitrile increased the ratio of reduced product to cross-linked product two-fold; a consquence, they argue, of acetonitrile's ability to compete with excited tyrosine for hydrogen atom donation to the uracilyl radical (itself a result of initial electron donation from excited tyrosine). Due to these results and the fact that tyrosine is a facile photoejector of an electron (Creed, 1984), Norris et al. contend it is the excitation of tyrosine and not BrU which leads to the photocoupled product. If this hypothesis is true, then thymidine should be capable of cross-linking to tyrosine; albeit with a smaller quantum yield as it has a lower

electron affinity and is not capable of bond homolysis as is BrU. Hicke et al. (1994) observed just such photocross-linking between *Oxytricha nova* telomeric DNA $(T_4G_4)_2$ and its associated telomere binding protein. In fact, Shetlar and coworkers (1992) observed cross-linking in model studies with thymidine much like that reported by Norris and coworkers. These researchers also argue that excitation of N-acetyltyrosine with subsequent electron ejection and capture by thymidine leads to radical coupling and adduct formation. As a result, nucleoprotein complexes should give improved cross-link yield if tyrosine (and not BrU as originally believed) can be selectively excited at a wavelength transparent to tryptophan.

Norris and coworkers measured the extinction coefficients of tryptophan and tyrosine derivatives at 308 and 325 nm. They discovered that while the extinction coefficient of the tryptophan derivative is three times that of the tyrosine derivative at 325 nm, it is 50 times that of tyrosine at 308 nm. As a result, they photocross-linked the BrU-RNA/MS-2 bacteriophage coat protein complex at 325 nm light and noted a cross-link yield of 66% after 10 h of irradiation as compared to a maximum cross-link yield of 40% at 308 nm light irradiation.

Taken together, the various investigations into BrdU photochemistry reviewed here indicate an extensively studied and potent aromatic amino acid cross-linking chromophrore. Despite the apparent incongruity between model studies and macromolecular systems, they all indicate that BrU based PhotoSELEX should be possible whether 308 or 325 nm light is used--especially if the protein target offers accessible aromatic residues. In fact, PhotoSELEX may converge fastest and yield the most resilient aptamers if 308 nm light is used. Once individual aptamers are identified, then maximum cross-link yield may result from irradiation at 325 nm.

The basic Fibroblast Growth Factor as a promising PhotoSELEX target

The basic Fibroblast Growth Factor (bFGF) is a 155-amino acid member of the Fibroblast Growth Factor family of polypeptides. These polypeptides are known regulators of cell proliferation, differentiation and motility (Baird and Klagsbrun, 1991; Vlodavsky et al., 1991), and they are all known to bind the polyanionic heparin molecule (Shing et al., 1984). While they were first identified as mitogens for fibroblasts, they are now known to play much more varied roles. In particular, FGFs are critical in normal physiological processes such as embryonic development (Baird and Klagsbrun, 1991), radiation-damaged DNA repair (Vladovsky et al., 1991) and wound healing. However, with the identification of oncogenes that encode proteins with 40-50% sequence homology to acidic FGF (aFGF) and bFGF, they have also been identified in pathological functions such as cancer angiogenesis and tumorigenesis. The following is a review of the biological and biochemical properties of bFGF relevant to its selection as the target for the PhotoSELEX project.

bFGF stimulates proliferation of an impressive array of cells. In particular, it is known to stimulate proliferation of all mesoderm-derived cells. It has also been shown to be mitogenic for neuroectodermal and ectodermal cells. However, as cells derived from the endoderm including thyroid, prostatic and pancreatic cells also respond to bFGF, it seems likely that bFGF is mitogenic for all tissues derived from these three embryonic germ layers (mesoderm, ectoderm and endoderm). In summary the list of cells stimulated to proliferate in response to binding bFGF include fibroblasts, endothelial cells, astrocytes, oligodendrocytes, neuroblasts, keratinocytes, osteoblasts, smooth muscle cells, melanocytes and megakaryocytes

(Godpodarowics, 1991). Another major function of bFGF is its ability to induce cell motility. While, as will be discussed later, this is essential to tissue repair through angiogenesis (translocation of endothelial cells), it can also be pathogenic as manifested during tumor metastisis. For example, bFGF is known to regulate motility of human hepatoma cells via an autocrine mechanism (Kin et al., 1997).

In addition to its mitogenic and chemotactic functions, bFGF also acts as a cell morphogen; stimulation of cell proliferation and differentiation are functions previously assumed to be mutually exclusive. Differentiation begins with induction of embryogenesis which is defined as "the developmentally significant interaction between closely associated but dissimilarly derived tissue masses" (Gospodarowicz, 1991). One consequence of this interaction is the selective binding of various morphogenic factors such as bFGF from the extracellular matrix (ECM) which subsequently stabililizes differentiation of embryonic tissues (Gospodaraowicz, 1991). It is the binding of bFGF with high-affinity cell-surface receptors which initiates all these functions.

High-affinity (K_d = 2-20 x 10^{-11} M) transmembrane receptors that posseses tyrosine kinase activity and which are phosphorylated upon binding mediate the cellular response to bFGF (Yayon et al., 1991). In addition to the high-affinity receptors, however, ECM and cell surface-bound heparan sulfate proteoglycans (HSPG) serve as a low-affinity reservoir for bFGF (Moscatelli, 1988; Vlodavsky et al., 1987). Yayon and coworkers (1991) were able to culture cells with only low-affinity binding, only high-affinity binding and both low plus high-affinity binding bFGF receptors. They observed that cells with low and with low plus high-affinity binding receptors bound bFGF while cells with only high-affinity receptors did not.

However, Yayon and coworkers also observed that when heparin was included in the binding medium for the HS-deficient cells, bFGF binding increased in a dose specific manner. This observation held even after incubation for 3 h in the absence of heparin. Based upon these results, Yayon et al. concluded that the ability of heparin to promote high-affinity binding is not a result of a protective function (Gospodarowics and Chen, 1986) but that it must induce a conformational change facilitating an induced-fit with the high-affinity receptor (1991).

However, Vlodavsky and coworkers (1991) recognized that while bFGF is found in nearly every kind of solid tissue and is a potent angiogenic factor capable of stimulating endothelial cell proliferation and motility, endothelial cell proliferation is normally quite low. In order to explain this apparent contradiction, and the fact that endothelial cell proliferation increases markedly in response to injury, they hypothesized that bFGF must be sequestered from its site of action (high-affinity receptors) by cell surface and ECM bound HSPG; this conclusion is very much in accord with Yayon and coworkers (1991). Only when released by heparin-like molecules, enzymes and intact cells are bFGF molecules presented to and bound by the high-affinity receptors (Vladovsky et al., 1991). No matter the mode of release, bFGF's putative heparin binding site is occupied. This is significant in view of the observation by Prestrelski and coworkers (1992) that bFGF undergoes a small but reproduceable conformation change upon binding heparin. In short, stringent regulation of bFGF's angiogenic function is crucial for maintenance of tissue health and wound repair.

Angiogenesis, a term first used in 1935 to describe newly forming blood vessels in the placenta (Hertig, 1935), is now used to describe the process by which

any tissue acquires a new and independent blood supply sufficient to facilitate growth (i.e. neovascularization). Folkman and others have elucidated the sequential steps in angiogenesis (Folkman and Klagsbrun, 1987). Initially, new capillaries begin from the sprouting of small venules. Angiogenic factors cause endothelial cells to translocate to the sprouting venules. These migrating endothelial cells align with one another to form a solid sprout about the venule. This is followed by lumen formation which results from endothelial cell curvature. As the sprout continues to grow, it joins with another nascent sprout to form a loop which is finally followed by blood flow.

In cancer, the hypothesis that tumor growth requires neovascularization was first explored in 1963 when Folkman et al. observed that tumors implanted in perfused organs failed to grow. When these tumors were subsequently implanted in mice, they became vascularized and grew (Folkman, 1970). Investigation by other researchers revealed that long-term perfused organs experience degredation of the capillary endothelium and therefore do not offer a suitable environment for neovascularization (Gimbrone et al., 1969). Based upon these results, Folkman concluded that the generation of solid tumors is indeed an angiogenesis-dependent event (1971).

In 1968, Greenblatt and Shubik showed that tumors could induce blood vessel growth even though they were separated from the vascular bed. This was the first evidence to support the notion that extracellular angiogenic agents must be responsible for neovascularization. Subsequently, several researchers reported the isolation of specific angiogenic factors (Folkman et al., 1971; Klagsbrun et al., 1976). The first angiogenic factor to be purified was bFGF (Gospodarowics et al., 1978).

When the primary amino acid sequence of bFGF became known (Esch et al., 1985; Abraham et al., 1986) as shown below, its origin was suddenly uncertain as it does not possess an obvious signal sequence although the cell surface receptors

```
                10                  20                  30                  40
MAAGS   ITTLP     ALPED    GGSGA    FPPGH    FKDPK    RLYCK    NGGFF
                50                  60                  70                  80
LRIHP   DGRVD    GVREK    SDPHI    KLQLQ    AEERG    VVSIK    GVCAN
                90                  100                 110                 120
RYLAM   KEDGR    LLASK    CVTDE    CFFFE    RLESN    NYNTY    RSRKY
                130                 140                 150
TSWYV   ALKRT    GQYKL    GSKTG    PGQKA    ILFLP     MSAKS
```

Figure I.4.1. human basic Fibroblast Growth Factor Sequence(155aa).

suggest an export mechanism, and it was clearly extracelluar. Mignatti et al. (1991,1992) answered the question of bFGF's origin when they conducted tests which illustrated that bFGF is in fact released from living cells, that it mediates the releasing cell's migration and that inhibitors of endo and exocytosis block its release while ER-Golgi inhibitors do not. Mechanisms suggested for the release of bFGF from the cell include release in response to heat shock (Jackson et al., 1992) and via an ABC-transporter pathway (Kuchler, 1993). While the mechanism of secretion is still unclear, it is clear that bFGF is an angiogenic factor for tumors.

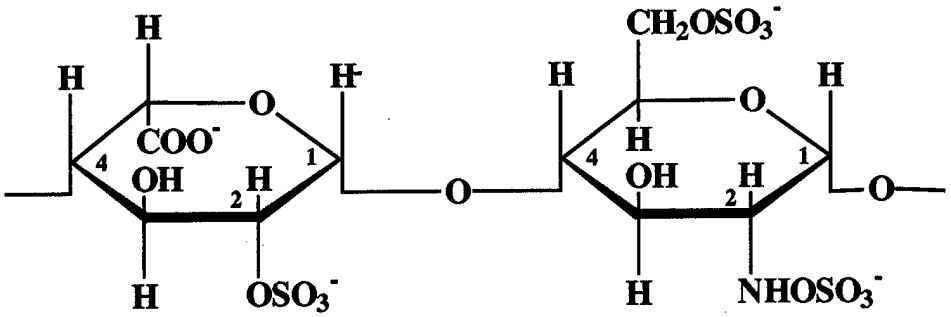
Nguyen and coworkers (1994) have observed elevated levels of bFGF in the urine and sera of patients with a wide array of cancers and thereby inferred that over expression of bFGF is prerequisite for tumor angiogenesis. These researchers also argued that such elevated bFGF levels should serve as an effective prognostic tool. In agreement with this conclusion, Nanus and coworkers (1993) and Dirix and coworkers (1997) observed a strong correlation between elevated bFGF levels and poor survival rates. However, other researchers have either not observed elevated

bFGF levels in cancer patients (Relf et al., 1997) or found unrelated factors which contributed to elevated levels of bFGF and thus have questioned bFGF level as a prognostic tool (Hsu et al., 1997). Despite the evident confusion about measured bFGF levels in patient body fluid, the observation that inhibition of tumor growth in nude mice by neutralizing bFGF antibodies puts its tumor angiognenic function on solid footing (Hori et al., 1991). With its potential as a therapeutic target and possible value as a diagnostic marker, the selection of bFGF as a target for the PhotoSELEX project seemed an interesting one. However, it is bFGF's ability to bind the polyanionic heparin molecule which makes it such an attractive target for the nascent PhotoSELEX technology.

Central to its rich and varied biochemical functions is the protection, control and activation bFGF gains from its binding to heparin (Gospodarowics and Cheng, 1986; Yayon et al., 1991). As mentioned at the outset, all members of the FGF family bind heparin and are alternatively known as heparin-binding growth factors (HBGF) (Burgess and Maciag, 1989). In view of its crucial role in the biochemistry of bFGF, and its relevance as a possible prototype to any ligand evolved by PhotoSELEX, a brief review of heparin is in order.

Heparin and heparan sulfate are glycosaminoglycans--unbranched polysaccharides of alternating α (1-4) linked N-glucosamine and L-iduronic acid carrying sulfate substituents in various regions (Kjellen and Lindahl, 1991). Heparin is initially formed as a polysaccharide chain with the structure [GlcA β1/4 GlcNac α1/4]n. This is subsequently N-deacylated/N-sulfated and finally undergoes C5 epimerization of GlcA to IdoA to yield the disaccharide repeating unit that appears in

Figure I.4.2 below.



L-Iduronate-2-sulfate N-Sulfo-D-glucosamine-6-sulfate

Heparin

Figure I.4.2. The heparin repeating disaccharide unit (Kjellen and Lindahl, 1991).

These modifications to the initial polymer result in a high degree of heterogeneity between heparin molecules. Despite this apparent diversity, heparin retains its high-affinity binding capacity to various proteins by conserving key moities. For instance, heparin's anti-coagulation properties result from its ability to bind anti-thrombin. Heparin's anti-thrombin binding region is a pentasaccharide incorporating several variants between individual heparin molecules. All though variable, all such pentasaccharides possess a key 3-O-sulfate group (Kjellen and Lindahl, 1991).

Heparin and heparan sulfate do not exist as free molecules *in vivo* but as constituents in proteoglycans. Proteoglycans (PG) are macromolecules composed of glycosoaminoglycan (GAG) side chains covalently bound to glycopeptide linkage regions. These glycopeptides are themselves covalently bound to a core protein through either the amide N of specific Asn residues (N-linked oligosaccharides) or the side chain O of either Ser or Thr from the core protein (O-linked oligosaccharides) (Voet and Voet, 1995). Heparin occurs almost exclusively in the

intracellular granules of most cells that line arterial walls. What's more, heparin has 2.5 sulfate resides per disaccharide unit making it the most polyanionic molecule found in mammalian tissue (Voet and Voet, 1995). Due to its polyanionic structure and its ability to bind FGFs, bFGF seemed a very suitable target for which PhotoSELEX might identify high quantum yield cross-linking/high affinity binding polyanionic ssDNA aptamers. Of course, this task should be rendered more tenable by a knowledge of bFGF's heparin binding site.

Identification of the heparin binding site in bFGF eluded researchers for many years but was successfully characterized in the early 1990s. The X-ray crystal structure of bFGF has been determined (Ericksson et al., 1991; Ericksson et al., 1993; Zhu et al., 1991; Zhang et al., 1991; Kastrup et al., 1997). (It should be noted that three main isoforms of bFGF have been examined (155aa, 154aa, and 146aa). These different isoforms apparently result from different sources and the correspondingly different methods of isolation and purification. Unfortunately, this has resulted in some ambiguity in numbering and identifying the amino acid residues in bFGF's primary structure. Most researchers begin numbering at the N-terminus of the isoform under their investigation without reference to other isoforms. To minimize confusion in this discussion and establish a consistent basis for comparing results of different researchers, two sets of numbers will identify each amino acid residue discussed. The first is that reported by the authors themselves. The second number in parentheses will indicate that residue's position within the 155aa isoform—that used in the present research.) Ericksson and co-workers (1993) determined that bFGF includes a Kunitz-type fold and 12 anti-parallel β-strands. They also identified a sulfate ion hydrogen-bonded to the side chains of Asn27(36), Arg120(129) and

Lys125(134) which they speculated identified the heparin binding site. This conclusion was substantiated by Thompson and coworkers (1994). These researchers used the position of such sulfate ions in X-ray crystal structures and pentasaccharide heparin-bFGF complex data to conduct site-directed mutagenesis studies. Based upon binding differences of various bFGF mutants to low molecular weight heparin as measured by isothermal titrating calorimetry and [NaCl] elution from heparin-Sepharose, they concluded that the heparin binding site included a discontinuous epitope involving Lys26(35), Asn27(36), Arg81(90), Lys119(128), Arg120(129), Thr121(130), Gln123(132), Lys125(134), Lys129(138), Gln134(143), and Lys135(144). Additional evidence corroborating this as the heparin binding site comes from the observed reduction in bFGF binding affinity for heparin upon reductive methylation studies on conserved Lys125(134) in aFGF (Harper and Lobb, 1988) and site-directed replacement of Lys125(134) with glutamic acid (Burgess et al., 1990).

Recently, Faham et al. (1996) successfully determined the X-ray structures of bFGF complexed with heparin-derived tetra and hexasaccharides at 1.9 and 2.2 angstroms, respectively. They reported that both molecules bound in similar regions of the bFGF surface containing residues Asn28(36), Arg121(129), Lys126(134) and Gln135(143); the hexasaccharide associated with an additional binding site formed by Lys27(35), Asn102(110), and Lys136(144). Also, Moy and coworkers (1996) determined the high-resolution solution structure of bFGF using multidimensional heteronuclear magnetic resonance. Again, their data were in agreement with the majority of X-ray crystallographic data acquired to date with one slight but interesting difference. They deduced that the overall structure of bFGF includes 11 extended

anti-parallel β-strands and a pseudo-3-fold axis of symmetry. In addition, they determined that one of the β-strands previously observed by X-ray diffraction is actually a helix-like conformation incorporating residues 131-136(131-136). As these sequences lie in the primary heparin binding site (128-138)(128-138), Moy and coworkers suggest that the helix-like conformation may be a crucial component to bFGF's recognition of heparin.

The importance of heparin and heparan sulfate binding to bFGF's biochemistry has been stated throughout this review. However, Moy and coworkers (1997) have recently elucidated the exact function played by heparin in facilitating bFGF binding to high-affinity transmembrane receptors. By employing NMR and dynamic light scattering techniques, they determined that the biologically active form of bFGF is a heparin-derived decasaccharide induced dimer or tetramer of bFGF. Finally, researchers have successfully characterized some of bFGF's individual chemical properties.

The chance for success of any SELEX protocol for a given protein is enhanced by a knowledge of its specific chemical behavior. Thompson and Fiddes (1991) helped provide such knowledge when they chemically characterized the cysteines of bFGF. There are four cysteine residues occurring at positions 34(34), 78(78), 96(96) and 101(101) in the primary amino acid structure of bFGF. The cysteines at 34(34) and 101(101) are completely conserved throughout the FGF family and thereby suggest the necessity of an intramolecular disulfide bond between these cysteines for biological activity. To assess the validity of this assumption, they employed site-specific mutagenesis and Ellman's reagent and discovered that Cys34(34) and Cys101(101) do not form an intramolecular disulfide bond in either

natural or recombinant bFGF, but that Cys78(78) and Cys96(96) do form intermolecular disulfide linkages resulting in multimeric structures in recombinant bFGF. Interestingly, they did not observe multimerization of natural bFGF even in the presence of $CuSO_4$. They discovered that Cys78(78) and Cys96(96) in natural bFGF formed disulfide linkages with glutathione (an artifact of purification) which prevents multimerization without any apparent permutation in structure.

It is clear from this discussion that bFGF plays many important biological functions from cell proliferation and differentiation to angiogenesis and the pathological function of tumorigenesis. Due to its biological significance and heparin-binding capacity, bFGF seemed a promising target for the PhotoSELEX protocol. Fortunately, bFGF is now a very well-characterized molecule and a good deal is known about its chemical properties. With this information in hand, the PhotoSELEX experiment was initiatied against the bFGF target.

CHAPTER II: EXPERIMENTAL PROCEDURES

Materials

Protein preparations. recombinant human basic Fibroblast Growth Factor$_{155}$ (bFGF) expressed in *E. coli*, was purchased from Bachem California (Torrence, CA), Lot Number SN 433. The lyophylized protein was resuspended in 1x phospate buffered saline (PBS) (Sambrook et al., 1989), 2 mM $MgCl_2$, and 0.01% human serum albumin (HSA) to a concentration of 10 μM. The resuspended protein was divided into 20 μL aliquots and stored at –20 °C until ready for use. Vascular Endothelial Growth Factor (VEGF) and Platelet Derived Growth Factor (PDGF) 1xPBS suspensions originally purchased from R&D Systems (Minneapolis, MN) were gratefully accepted from Louis Green and Rob Jenison, respectfully, of NexStar Pharmaceuticals, Inc (Boulder, CO). Human blood serum was gratefully accepted from Barbara McBride and Andrew Stephens, also of NexStar Pharmaceuticals.

Oligonucleotide preparations. The completely random ssDNA library (zero rounds of affinity selection (0xA) and zero rounds of covalent selection (0xP), was prepared at NexStar Pharmaceuticals with an Applied Biosystems Model 394 solid phase oligonucleotide synthesizer employing established protocols (5'-GGGAGGACGATG-30n-CAGACGACGAGCGGGA-3'). In this prototype sequence 30n indicates the 30 nucleotide random region. The ssDNA library evolved from five rounds of affinity selection (5xA/0xP) for the bFGF target was generated and provided by Brian Collins and Mike Willis of NexStar Pharmaceuticals. The 5' and 3' primers required for polymerase chain reaction (PCR) amplification of these libraries during rounds of PhotoSELEX were also prepared at NexStar

Pharmaceuticals on the Applied Biosystems Model 394 oligonucleotide synthesizer: 5' primer (GGGAGGACGATGCGG) and 3' primer (biotin-sp-biotin-sp-TCCCGCTCGTCGTCTG). All phosphoramidites required for solid phase synthesis including biotin-sp-biotin-sp-thymidine were purchased from Glen Research (Sterling, VA). Unmodified deoxynucleoside triphosphates (dATP, dGTP, and dCTP) were purchased as 10 mM glycerol suspensions from Pharmacia Biotech (Piscataway, NJ). The modifed deoxynucleoside triphosphate 5-bromo-2'-deoxynucleoside-5'-triphosphate was purchased from Sigma Chemical Co. (St. Louis, MO) as a lyophylized sodium salt which was resuspended in deionized water to 10 mM. [γ-^{32}P] ATP used in T4 polyneucleotide kinase 5' labeling of ssDNA was acquired from NEN Life Science Products (Boston, MA).

Chemicals and reagents. Dimethyl sulfate, formic acid, and iron(III) chloride required to complete Maxam-Gilbert sequencing reactions were obtained from Aldrich Chemical Co. (Milwaukee, WI). Anhydrous hydrazine, dimethyl sulfate buffer, yeast tRNA, and calf thymus DNA also required for Maxam-Gilbert sequencing were purchased from Sigma Chemical Co. Snake venom phosphodiesterase I and calf alkaline phosphatase required for digestion of oligonucleotide in preparation for electrospray mass spectrometry (ESMS) were purchased from Pharmacia Biotech. Trypsin required for digestion of bFGF in preparation for Edman degredation and ESMS was purchased from Sigma Chemical Co.

Methods

The PhotoSELEX experiments. The fundamental and conceptual aspects of a round of PhotoSELEX are described in the Introduction. Furthermore, the

photocross-linking reaction conditions for each PhotoSELEX round (including concentration of DNA, concentration of bFGF, and time of irradiation (specified as delivered pulses)) are recorded in Chapter III: The PhotoSELEX experiments. The following is a detailed description of all other steps required to complete a PhotoSELEX round.

PCR amplification in which 5-BrdUTP has been substituted for TTP of 15 pmol ssDNA was performed to generate sequences incorporating the chromophore for both the 0xA/0xP and 5xA/0xP libraries (#xA ≡ # of rounds of affinity SELEX through which a library has evolved. #xP ≡ # of rounds of PhotoSELEX through which a library has evolved). The PCR reaction for these noncross-linked libraries employed a 1 mL reaction volume with a 1:1 mixture (12.5 units of each) of *Taq* polymerase enzyme (Perkin Elmer Inc., Norwalk, CT) and *Pwo* polymerase enzyme (Boehringer-Mannheim Inc., Indianapolis, IN) in a standard *Taq* PCR buffer including 15 mM $MgCl_2$. To maximize amplification with the modified dNTP, all reagents other than polymerase enzymes were subjected to a Hot Start (incubation at 75 °C for 3 min prior to amplification). In these reactions, 13 cycles of amplification {1 cycle = 1) 95 °C/40 s, 2) 55 °C/60 s, 3) 72 °C/120 s} with an MJ Research Minicycler (Watertown, MA) were used. The resulting dsDNA was concentrated with Centri-Sep 30k MWCO filters (Schlinker and Schuell Inc., Keene, IN) by centrifugation at 12,000 rpm for 10 min in an Eppendorf Centrifuge (#5415C). Resuspension of concentrated DNA from the filter surface was accomplished with 45 μL of standard formamide gel loading buffer (Sambrook et al., 1989) including 0.1% sodium dodecyl sulfate (SDS). To purify and isolate the recovered dsDNA to obtain sense ssDNA, the dsDNA was

heated at 95 °C for 2 min and segregated on a 12% polyacrylamide gel incorporating 7 M urea. The migration of the anti-sense strand was slowed due to the presence of two biotin groups covalently bound to the 3' primer, thus permitting ready segregation.

Sense ssDNA was located on the polyacrylamide gel by UV shadowing and subsequently excised. This gel band was crushed in 300 μL of 2 mM EDTA using a 1 cc syringe plunger until a homogeneous slurry was obtained. The slurry was vortexed at room temperature from 3 to 8 h and then centrifuged through a 0.45 μm cellulose acetate Costar filter (Corning Inc., Corning, NY) at 14,000 rpm/10 min. ssDNA was finally concentrated by standard NaOAc/ethanol precipitation and quantified by UV spectroscopy. To facilitate detection and quantification, 2 pmol each of the recovered ssDNA libraries were then 5'-radiolabelled with T4 polynucleotide kinase (PNK) (Boehringer-Mannheim) and purified on a 14% polyacrylamide gel. The position of the radiolabeled DNA was determined with 1.5 min exposure of Kodak X-OMAT AR film (Eastman Kodak Co., Rochester, NY). The radiolabelled band was excised and recovered as described above. These libraries were then subjected to the photocross-linking reaction with bFGF.

The rationale for selection of ssDNA and bFGF concentrations used for photocross-linking reactions in each round of PhotoSELEX are described in Chapter III, as previously mentioned. Once the DNA/protein concentrations had been selected, 200 μL of 2x DNA was doped with 2,000 cpm/μL 5'-^{32}P-labelled DNA. This was combined with 200 μL of 2x bFGF and incubated at 37 °C for 15 min. Dilution to desired 2x concentrations was accomplished with a cross-link binding buffer (XLBB) (1x PBS/2 mM $MgCl_2$/0.01% HSA/1.0 mM dithiothreitol (DTT)). The mixture was

then placed in a 1.5 mL semi-micron methacrylate cuvette possessing a 1 cm path length (Fisher Scientific Co., Pittsburgh, PA) which was subsequently placed in a 37 °C cuvette holder. The cuvette was then irradiated with 308 nm light generated by a Lumonix Model EX 700 XeCl excimer laser. Time of irradiation was specified by the number of 20 ns pulses delivered to the sample. The laser was set to deliver 20 such pulses per second with a total pulse energy of 175 mJ. Each laser pulse was attenuated by passing through a quartz convex lens with a focal length of 10 cm. In each reaction, the sample was placed 40 cm beyond the focal point significantly reducing the number of incident photons. With this configuration, the sample received approximately 11 mJ/pulse of 308 nm light. Photocross-linking reactions used for screening, maximizing cross-link yield, and for evaluating diagnostic potential were also conducted in this way.

Partitioning of sequences capable of photocross-linking with bFGF from those incapable of such cross-linking was accomplished by PAGE as mentioned in the Introduction. This was done by first reducing the cross-linking reaction volume by centrifugation through 30k MWCO Centri-Sep filters at 12,000 rpm/15 min. Cross-linked oligonucleotide/protein adducts, noncross-linked ssDNA, unreacted bFGF, and photo-damaged ssDNA were all recovered from the filter by washing with the same formamide gel loading buffer described above. This was heated at 95 °C for 2 min and then partitioned by 12% PAGE with 7 M urea. Identification of cross-linked material was accomplished by 15 h exposure of Kodak BioMax MR film at –70 °C to create an autoradiogram. As $bFGF_{155}$ and 61-nt ssDNA sequences are both ≈ 18 kDa, cross-linked DNA migrated in the gel about half as fast as free DNA. The cross-linked

material was excised from the gel and then crushed to a slurry in 400 μL proteinase K buffer (0.1 M Tris/pH 7.5, 0.05 M NaCl, 0.02 M EDTA, 0.5% SDS).

To remove the protein portion, the cross-linked adduct was subjected to proteinase K digestion. This was accomplished by adding 100 μL of 4 mg/mL proteinase K in buffer to each 400 μL gel slurry containing the cross-linked adduct and incubated at 48 °C for 75 min. To promote digestion, 50 μL of 7 M urea was then added and incubating for another 15 min at 48 °C. After this period of incubation, two more additions of 100 μL of 7 M urea and 15 min incubations at 48 °C were completed. Finally, the digested adduct was extracted with phenol/chloroform to remove enzyme and amino acids from the oligonucleotides and then concentrated by NaOAc/ethanol precipitation. This completed one round of PhotoSELEX. Prior to initiation of each subsequent round of PhotoSELEX, a pilot PCR using 5'-^{32}P-radiolabelled primer was performed to identify the number of cycles which would give maximum PCR amplification (typically 16-23 cycles). As described in the PhotoSELEX experiments chapter, a pilot PCR is a microscale reaction (e.g. 1:10) in which aliquots are removed at predetermined intervals to monitor the extent of amplification. Half of the recovered oligonucleotide sample from the previous round was then used for PCR amplification to generate the new library. Monitoring PhotoSELEX evolution was accomplished by measuring photocross-link yield which is described below in Photocross-link yield and binding affinity screening. Evolution was also monitored through measurement of binding affinities of the evolved libraries which was accomplished by nitrocellulose filter binding also described below.

Cloning and sequencing of the PhotoSELEX libraries. Identification of

individual sequences began with PCR amplification of the three evolved libraries with new primers which would facilitate incorporation into plasmid DNA. After purification by phenol/chloroform extraction and NaOAc/ethanol precipitation, the amplified sequences were inserted into pUC 18 circular DNA plasmid with Hind III and Eco RI restriction enzymes. After purification, the inserts were ligated with T4 DNA ligase. XL-1 Blue ultracompetent cells were then transformed with the plasmid DNA containing sequence inserts by 30 min/4 °C incubation followed by 45 s/42 °C heat pulse. After incubation of the transformed cells in SOC medium at 37 °C for 1 h, the cells were plated onto an agar plate containing X-gal, isopropyl-thiogalactoside (IPTG) and ampicilin. Colonies were permitted to grow overnight at room temperature. Individual colonies were identified and cultured in 5 mL of SOC medium.

Plasmids were purified with the PERFECTprep Plasmid DNA kit, 5 prime→ 3 prime, Inc. (Boulder, CO). This involved the following: 1)1500 µL of selected cultures were extracted, centrifuged and freeze-thawed to predispose the cells to lysing, 2)100 µL of 50 mM Tris-HCl/pH 7.6, 10 mM EDTA/pH 8.0, 100 µg/mL RNase and 7.75 mg/mL lysozyme were added followed by 100µL of 0.2 N NaOH, and 1.0% SDS, 3)after 10 min, 100 µL of 1.32 M KOAc/pH 5.2 was added to each preparation and the mixture centrifuged. Bacterial contaminants were then removed by Perfect Prep Spin Columns containing a DNA binding matrix. Bound DNA was subsequently removed by incubation with TNE at 50 °C.

Finally, the individual sequences of cloned aptamers were determined using the

Dye Terminator Cycle Sequencing kit from Perkin-Elmer. This required PCR amplification of the plasmids which generated flourescent-labelled sequences. Sequence ladders were then read by the laboratory of Dr. Brian Kotzin, National Jewish Hospital, Denver, CO.

Photocross-link yield and binding affinity screening. Screening for photocross-linking quantum yield was completed by first radiolabeling 2 pmol of selected individual sequences with T4 PNK and [γ-^{32}P-ATP]. Once labeled, 100 µL of 2,000 cpm/µL 5'-^{32}P-labeled sequences (2x DNA) ($\approx$ 2 fmol) was combined with 100 µL of 50 nM bFGF in 1x cross-link binding buffer (XLBB) to give a 1x photocross-linking mixture of 1,000 cpm/µL ssDNA/25 nM bFGF. After incubation at 37 °C for 15 min, each reaction mixture was irradiated with 1,000 pulses of 308 nm light as described above and a 20 µL aliquot removed. Each 20 µL aliquot was treated by adding 20 µL of formamide gel buffer, heating at 95 °C for 2 min, and loading the entire volume on a 12% PAGE incorporating 7 M urea. The cross-link yield was then assessed by exposing the gel to a phosphor-image plate for 30 min with subsequent radiation counting accomplished with Fujix MacBas ver. 2.0 software (FujiPhoto Film Co., Stamford, CT). All radiation in the entire lane was counted including photodegraded oligonucleotide to determine total counts in a lane. Only the small fragment appearing in the precise cross-link band was credited as cross-link. This rigorous counting method was also employed when assessing evolution of libraries in the PhotoSELEX experiments.

Nitrocellulose filter binding was used to assess aptamer (as well as library) affinity for bFGF. This was accomplished by first preparing 2,000 cpm/30 µL 5'-^{32}P-

labeled material (≈ 2-5 fmol) from each library or chosen sequence (2x DNA). Each 2x oligonucleotide sample was mixed with 30 μL of each bFGF concentration to be evaluated in a 96 well tray. bFGF concentrations were prepared by serial dilution using 1x XLBB. After incubation at 37 °C for 15 min, a multi-pipeter was used to load 40 μL of each DNA/bFGF mixture onto a pre-washed Gibco nitrocellulose binding curve apparatus followed by washing with a dissociation wash buffer. Bound DNA was then measured by phosphor-image and a binding curve for the percent bound DNA vs. [bFGF] calculated by Kaleidograph ver. 3.09 software (Synergy Software Co., Reading, PA).

Evaluation of the diagnostic potential of 06.15 and 06.50. Two sequences identified by screening were used to evaluate the potential of PhotoSELEX to identify diagnostic tools. Photocross-linking reactions were completed with sequences identified as 06.15 and 06.50 in which 50 μL of 2,000 cpm/μL ssDNA (2x DNA) was combined with 50 μL of sequentially diluted bFGF concentrations. Each cross-link reaction was performed as described above by irradiating each mixture with 1,500 pulses of 308 nm light. As before, 20 μL aliquots of each reaction were removed. In this case, however, each aliquot was mixed with 20 μL of 2x SDS gel loading buffer (0.1 M Tris-Cl/pH 6.8, 0.2 M DTT, 4% SDS, 2x dye (40% glycerol, 0.4% bromophenol blue, 60% water)). These mixtures were heated at 95 °C for 3 min and loaded on a 5% SDS-Stacking PAGE/12% SDS-Resolving PAGE. SDS-PAGE preparation including SDS-PAGE buffer is described by Sambrook and coworkers (1989). The recipe for preparation of the gel observed the following: 30 mL of 12%

SDS-Resolving PAGE (9.9 mL H_2O, 12.0 mL 30% acrylamide, 7.5 mL 1.5 M Tris (pH 6.8), 0.3 mL 10% SDS, 0.3 mL ammonium persulfate, 0.012 mL TEMED); 6 mL of 5% SDS-Stacking PAGE (4.1 mL H_2O, 1.0 mL 30% acrylamide, 0.75 mL 1.0 M Tris (pH 6.8). Resulting cross-link yield was assessed by phosphor-imaging as previously described.

Edman degradation and electrospray mass spectrometric determination of the bFGF cross-linking amino acid. Preparation of cross-linked oligonucleotide/protein adduct is briefly described in the corresponding section in the PhotoSELEX experiments chapter; it is prerequisite for both Edman degradation and ESMS sequencing. To generate sufficient quantities of sequences 06.15 and 06.50 for these reactions, 30 pmol of each sequence prepared by solid phase synthesis was used as a template in a 1.5 mL PCR amplification. Pilot PCR revealed maximum amplification at 16-19 cycles using the PCR protocol described previously. Purification, quantitation and radiolabeling were also accomplished as described above. Once these sequences were quantified by UV spectroscopy, 500 µL of 1.70 µM/1,000 cpm/µL 5'-^{32}P-lableled samples of both sequences were prepared with XLBB (2x DNA). Two-500 µL samples of 0.567 µM bFGF were also prepared (2x bFGF) and the solutions mixed to obtain 1x cross-linking solutions. Both systems were irradiated with 1,500 pulses of 308 nm light after 15 min of 37 °C incubation as previously described. After photocross-linking, 400 µL of 5 mg/mL trypsin suspended in 0.1 M Tris/pH 8.0 was added to each reaction volume and incubated overnight at 37 °C. Each system was then extracted with phenol/chloroform, precipitated with NaOAc/ethanol, and resusupended in 30 µL H_2O/30 µL formamide gel loading buffer. After heating at

95 °C, the trypsin-digested oligonucleotide/peptide adducts were partitioned on a 40 cm 14% PAGE incorporating 7 M urea. Bands migrating slower than free DNA were excised for both 06.15 and 06.50 sequences and crushed to a homogeneous slurry in 400 μL of de-ionized H_2O. Cross-linked adduct was recovered by centrifugation of the slurry through 0.45 μm Costar filters (14,000 rpm/15 min) followed by NaOAc/ethanol precipitation and resuspension in 100 μL de-ionized water. The amount of recovered trypsin-digested/cross-linked adduct was estimated by measuring the activity of 1 μL of this solution in 3 mL of scintillation fluid. For example, recovered 06.50/bFGF trypsin-digested cross-linked adduct (06.50 XL-dig) possessed 544 cpm/μL activity with a total volume of 100 μL which gave 54,400 cpm total recovered activity. However, nearly one week had elapsed since the initial photocross-linking reaction had been performed. Therefore, decay was accounted for by assessing activity of original 5'-^{32}P-labeled 06.50 by measuring activity of the same volume as originally doped into cross-linking reaction. Finally, the fraction recovered was multiplied by the total quantity of 06.50 present in the reaction to give the amount 06.50 XL-dig recovered: (54,400 cpm/304,385 cpm) x 833.3 pmol = 152 pmol. Approximately 60 pmol of 06.50 XL-dig and approximately 30 pmol of all three 06.15 XL-dig bands (T,M,B) were then reprecipitated. Since alkylation of thiol residues like cysteine permits more sensitive detection by Edman degredation sequencing, all samples were reduced with DTT and alkylated with iodoacetamide. This was precautionary as it was unknown *a priori* whether the isolated peptide fragments possessed thiol residues. All samples were finally reprecipitated by NaOAc/ethanol and resuspended in 40 μL of de-ionized water. These samples were then sent to Dr.

Dave McCourt at Midwest Analytical, Inc., St Louis, MO, for automated Edman degredation sequencing.

The preparation of 06.50 XL-dig for analysis by ESMS required digestion of the oligonucleotide fragment. Lyophylized snake venom phosphodiesterase I (SVP) was resuspended to a concentration of 1 mg/mL with dilution buffer (110 mM Tris/pH 8.9, 110 mM NaCl, 15 mM $MgCl_2$, and 50% glycerol in H_2O). Digestion of the oligonucleotide component of the 60 pmol sample of 06.50 XL-dig was accomplished by: 1) dilution to a volume of 43.3 μL with de-ionized H_2O, 2) addition of 0.8 μL 1 M $MgCl_2$, 3) addition of 3.5 μL 0.5 M Tris/pH 7.5, 4) addition of 2.4 μL of resuspended SVP to give a total reaction volume of 50 μL. The reaction mixture was then incubated overnight at 37 °C. The SVP enzyme was deactivated by heating the mixture at 95 °C for 2 min. Dephosphorylation was then accomplished by first completing a 1:20 dilution of concentrated alkaline phosphatase (AP) with de-ionized H_2O. Next, 5.8 μL of 10x One Phor All Buffer™ and 2 μL of the diluted alkaline phosphatase were added and the mixture incubated at 37 °C for 30 min followed by heating at 95 °C for 2 min to deactivate the enzyme. As a control, 60 pmol of 06.50 (without cross-link) was subjected to identical digestion conditions. Both samples were then analyzed by ESMS.

Dr. Katheryn Resing, Department of Chemistry and Biochemistry at the University of Colorado, Boulder, completed the mass spectral analysis of 06.50 XL-dig. Mass spectrometry (MS) and collision-induced dissociation (MS/MS) were accomplished with an API-III$^+$ triple-quadropole mass spectrometer (Sciex) equipped

with a nebulization assisted electrospray ion (ESI) source and a high pressure collision cell. Samples were analyzed by HPLC using a 500 μm i.d. column packed manually with 300 Å Columbus C18 resin (Phenomenix, Torrence, CA), equilibrated with 0.1% formic acid eluting at 40 μL/min and directly coupled to the mass spectrometer electrospray interface. The peptide-oligonucleotide cross-link was eluted from the column with a 2%/min gradient with acetonitrile at 20 μL/min, with the orifice voltage at 75 V, scanning at 0.6 s/scan. The 60 pmol of peptide/dinucleotide was used for two injections: the first 20 pmol for mass scans from m/z 300-1600 g/mol with elution from 0-40% acetonitrile for determining the retention time and the second 40 pmol for sequencing the peptide-dinucleotide cross-link. The identification of the HPLC peak associated with the cross-linked peptide was facilitated by comparison of the chromatogram with that resulting from injection of the control bearing compounds (SVP/AP digested/noncross-linked 06.50).

Maxam-Gilbert determination of the cross-linking position of 06.50.

Sequencing was initiated by radiolabelling an estimated 2 pmol of 06.50 XL-dig with T4 PNK and [γ-^{32}P-ATP] with subsequent PAGE purification as previously described. To ensure minimal salt content in the radiolabelled sample, 2x NaOAc/ethanol precipitations were completed. The activity of the labeled cross-linked/digested adduct was determined by Cerenkov counting and diluted to 7,500 cpm/μL Finally, the traditional method of Maxam-Gilbert sequencing was completed precisely as described by Sambrook and coworkers (1989). The resulting oligonucleotide digest reactions were then analyzed by 10% PAGE incorporating 7 M urea.

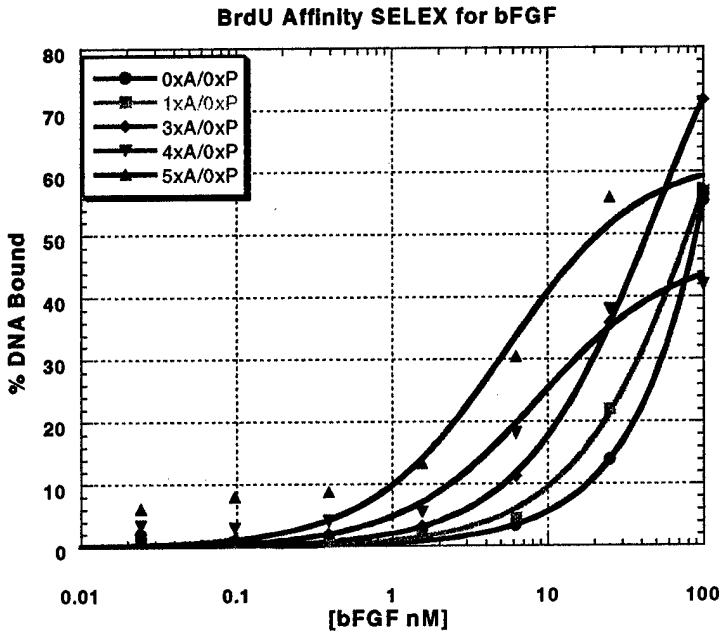
CHAPTER III: THE PHOTOSELEX EXPERIMENTS

The fundamental objective of this research was to explore the potential for PhotoSELEX to identify high-affinity binding/high-quantum yield photocross-linking oligonucleotide aptamers for the bFGF target. A significant aspect of this exploration was a detailed and systematic investigation of the experimental parameters required for PhotoSELEX feasibility. This chapter describes the results of this investigation and the essential conclusions drawn regarding the associated potential of PhotoSELEX as a high through-put identifier of therapeutic and diagnostic aptamers.

As mentioned in the Introduction, the hypothesized power of PhotoSELEX to rapidly evolve target solutions lies partly in the precise 'lock and key' fit required between aptamer and target to permit photoinduced bond formation. In addition, the high resolving capacity of PAGE which is used to segregate cross-linking aptamers from noncross-linking aptamers should permit very efficient partitioning. Since PhotoSELEX generates covalently bound oligonucleotide/protein adducts, the much more stringent conditions of denaturing PAGE can be used to reduce background relative to the mild conditions of nitrocellolose partitioning employed in affinity SELEX. Unfortunately, it is just these pressures to converge that could break the PhotoSELEX protocol if it is pursued too aggressively. For example, PhotoSELEX may not be able to adequately sample all sequence space with the subsequent loss of winning aptamers if the initial degree of randomization is too great. Therefore, the first experiments on optimizing the PhotoSELEX protocol were initiated on a library which had evolved through five rounds of standard affinity selection for the bFGF target.

Five rounds of affinity selection (5xA) substantially reduces the degree of

randomization for a target relative to the completely unevolved library (0xA). In the case of bFGF, affinity selection of the BrdU substituted DNA library improves the K_d from 200 nM for the completely random library to 5.3 nM for the fifth round affinity library as shown in Figure R.1.1 below. While this seemed a reasonable starting point for the nascent PhotoSELEX protocol, the DNA/protein ratio still had to be optimized in the photocross-linking reactions for maximum cross-linking aptamer recovery/evolution.



Library	0xA/0xP	1xA/0xP	3xA/0xP	4xA/0xP	5xA/0xP
K_d	200 nM	124 nM	51 nM	8.6 nM	5.3 nM

Figure R.1.1. Evolution of BrdU affinity SELEX for the bFGF target as measured by library affinity.

Irvine and coworkers (1991) have developed quantitative measures for facilitating rapid convergence of the affinity SELEX protocol. They determined that

such factors as the original library K_d for the target, the efficiency of partitioning, and background nonspecific binding among others were critical factors influencing rapid and successful convergence. Furthermore, they argue that while a high protein concentration relative to DNA permits ready capture of high-affinity aptamers even when they are rare in the initial library, maintaining such a ratio throughout the SELEX will drastically increase the time to convergence. Therefore, Irvine and coworkers suggest using relatively high protein concentrations in the first rounds of affinity SELEX to capture all the best binding aptamers followed by increasing protein stringency which will facilitate rapid enrichment of these high-affinity binders. A guess as to the initial protein concentration can be determined using the following from Irvine et al., (1991):

$$T_{no} = \{(K_d) + (\text{oligo})\}(0.7)(BG/CP)^{2/3}$$

where

$T_{no} \equiv$ The near optimal protein concentration. This is the concentration which should permit most efficient capture of all desired aptamers.

$K_d \equiv$ The library dissociation constant for the target protein. This is a measure of the library's affinity for the target.

$BG \equiv$ The nonspecific background. This is a measure of the fraction of the library which partitions as background noise. It reflects the facility of the partitioning mechanism.

oligo $\equiv$ The oligonucleotide concentration. This is the concentration of the oligonucleotide library under which the SELEX will be performed.

$CP \equiv$ The specific signal binding. This is a measure of the partitioning of desired aptamers due to binding or cross-linking.

Starting with a library K_d of 5.3 nM, taking 50 nM as a standard DNA concentration and assuming a partition efficiency (initially) of 10 to 15%, T_{no} = 10 nM. Therefore, the first round of PhotoSELEX utilized a DNA/protein ratio of 50 nM (5xA/0xP):10

nM (bFGF) where 0xP indicates zero rounds of PhotoSELEX.

The other selective pressure to determine in PhotoSELEX is the total number of photons to be delivered to the DNA/protein system; fewer photons delivered applies greater photochemical stringency. Collins et al., (manuscript in preparation) were able to identify a photocross-linking ligand to bFGF via post-affinity SELEX modification. This technique employs standard affinity SELEX followed by cloning and sequencing of the evolved pool. The resulting sequences are then modified by replacing T with BrdU wherever it appears in the random region. Subsequent screening by Collins et al., then revealed a sequence capable of high-affinity binding and moderate cross-linking ability. These researchers determined this aptamer (from here on referred to as 225-61) gave 15% cross-link yield after 1500 pulses of 308 nm light at protein excess. With this as a reference, it was determinined that the initial PhotoSELEX photocross-linking reaction would be completed at 250 pulses of 308 nm light. This established all the parameters for the first PhotoSELEX round: 50 nM DNA/10 nM bFGF/250 pulses of 308 nm light. After the photocross-linking reaction was completed, the cross-linking aptamers were partitioned from the noncross-linking aptamers by 12% PAGE incorporating 7 M urea with formamide gel loading buffer. Next, the protein component of the adduct was digested in preparation for PCR amplification.

The ability to accurately replicate DNA (fidelity) is as critical to the success of PhotoSELEX as it is to the progeny of living organisms. Mutations (especially at the site of cross-linking) will generally be deleterious and may prohibit the survival of a given cross-linking aptamer through a PhotoSELEX round. In order to minimize the potential for PCR mutations or abortions resulting from large protein fragments

covalently bound to the oligonucleotide templates, Proteinase K was chosen to digest the bFGF molecule from the cross-linked adduct. Proteinase K is an extremely active subtilisin-related serine endopeptidase capable of cleaving peptides in an X↓Y fashion where X = aliphatic, aromatic, or hydrophobic amino acids and Y = any amino acid (Boehringer-Mannheim, 1998). Unfortunately, even complete digestion of a bFGF molecule from a hypothesized cross-linking aptamer will still yield a perturbed oligonucleotide template as shown in Figure R.1.2 below.

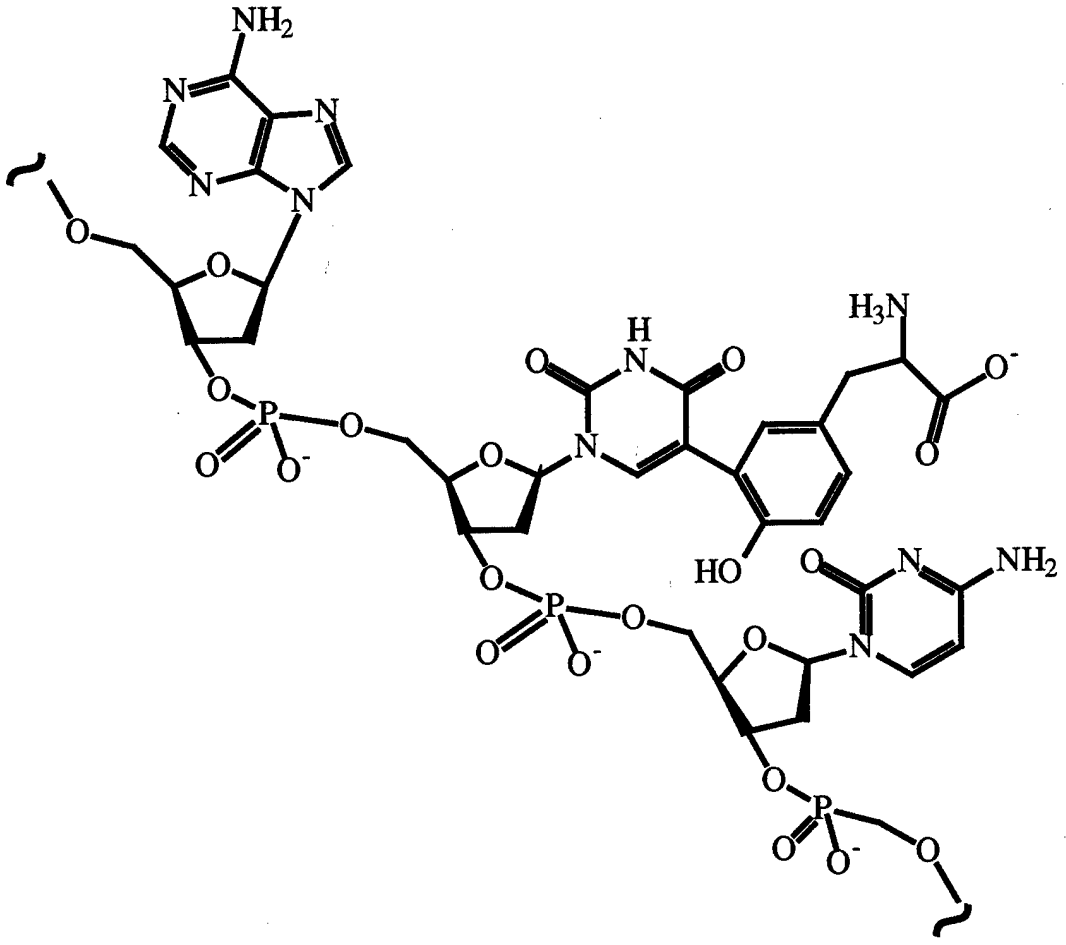


Figure R.1.2. Oligonucleotide section with putative cross-link to a tyrosine residue (shown in red) of bFGF after complete digestion with Proteinase K.

The first amplification of digested, cross-linking oligonucleotide sequences resulting from a PhotoSELEX round (5xA/0xP XL-dig) was completed with the thermostable polymerase enzyme *Pyrococcus woesei* (*Pwo*). *Pwo* is an extremely processive 5'-3' DNA polymerase possessing 3'-5' exonuclease activity and has been shown to give better PCR yield than *thermos aquaticus (Taq)*. For instance, typical yield from a 1000 μL *Pwo*-PCR reaction using 15 pmol of DNA template and dNTPs where TTP is replaced with BrdUTP is 250 pmol of PAGE-purified DNA. Conversely, an identical *Taq*-PCR reaction typically yields no better than 150 pmol of

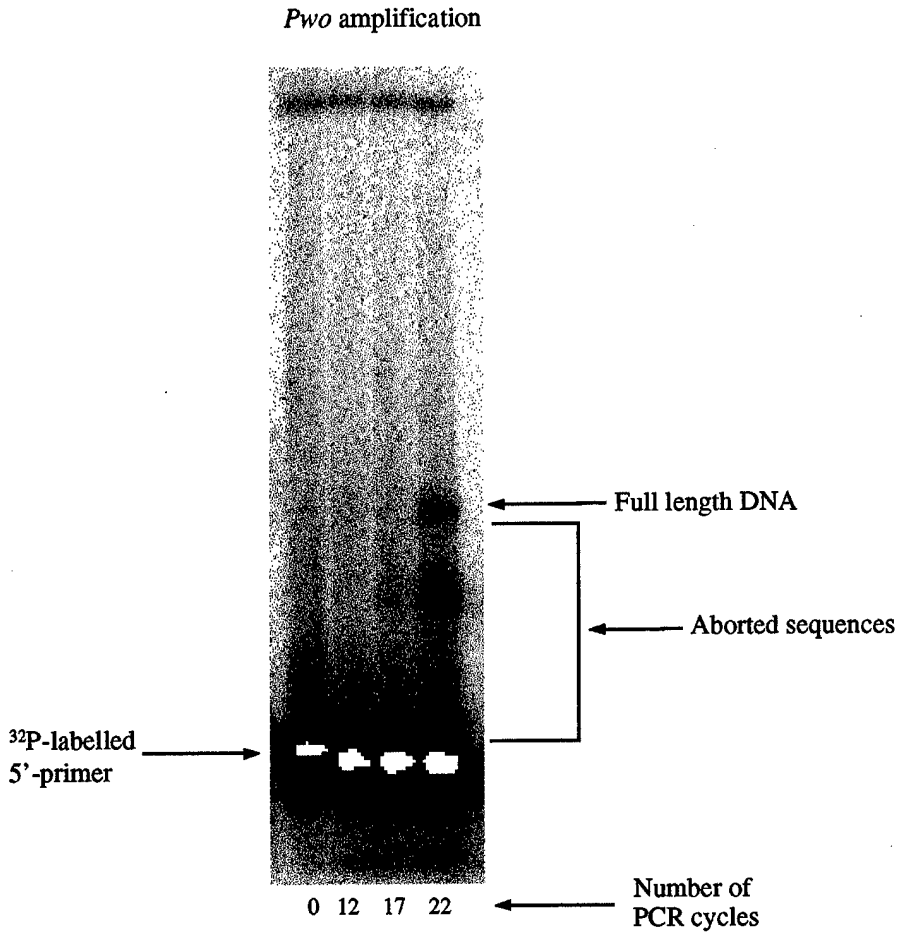


Figure R.1.3. PAGE of pilot PCR amplification of 5xA/0xP XL-dig using *Pwo* DNA polymerase.

DNA. Additionally, *Pwo* exhibits ten-fold greater fidelity in DNA synthesis relative to *Taq* (Boehringer-Mannheim, 1998). However, amplification of the digested/cross-linking aptamers resulting from PAGE partitioning of the first photocross-linking reaction by *Pwo* resulted in a greater than ten-fold reduction of typical yield. Moreover, Figure R.1.3 above illustrates the PAGE analysis of a pilot PCR amplification of digested cross-linking aptamers with 5'-^{32}P-radiolabelled 5' primer. A pilot PCR is simply a microscale (e.g. 1:10) PCR amplification; it is used to identify the number of PCR cycles which give maximum amplification for the reaction conditions. These reaction conditions yielded only a very faint product band and numerous abortions even after 22 cycles of PCR amplification.

When the DNA from the full-scale amplification was recovered, ^{32}P-radiolabelled at the 5' end with T4 polynucleotide kinase, and cross-linked again with bFGF, all cross-link signal originally present had been lost as shown in Figure R.1.4 on the following page. Evidently, *Pwo*'s tenacious processivity essentially renders it incapable of amplifying past the bFGF fragment remaining after Proteinase K digestion of the cross-link adduct. It is probable that the small amount of DNA recovered from the *Pwo*-PCR amplification is either background or highly-mutated versions of cross-linking aptamers. In either case, it was clear that PhotoSELEX would not proceed past one round much less converge with *Pwo*-PCR amplification of cross-link/digest sequences.

While *Taq* is also considered a fairly processive polymerase enzyme, its yield in PCR amplifications indicated it may not be as processive as *Pwo* (*vida infra*). It therefore seemed possible that a PCR reaction which benefited from *Taq*'s somewhat looser binding to the DNA template (and thus a potential ability to amplify past the

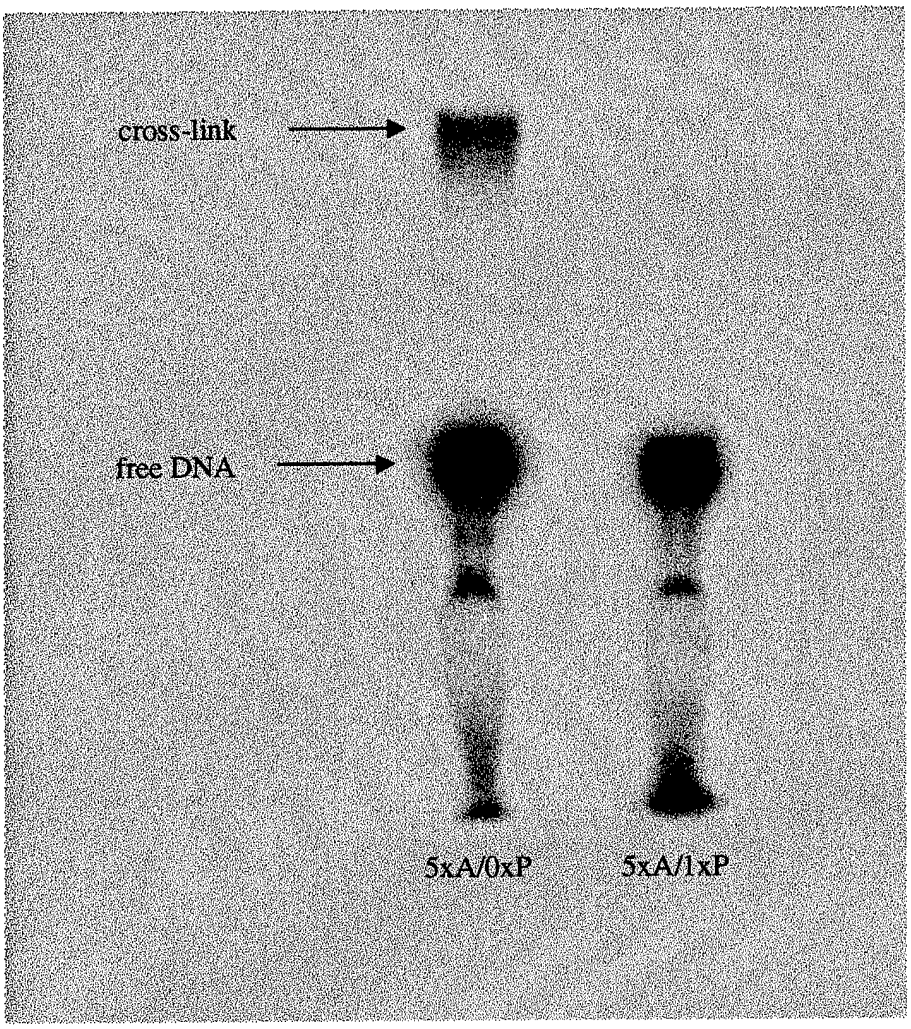


Figure R.1.4. Photocross-link reaction of 5xA/0xP and 5xA/1xP (*Pwo*) at 50 nM DNA/10 nM bFGF/250 pulses 308 nm light.

cross-link fragment) and *Pwo*'s high-fidelity, high-yield amplification would permit PhotoSELEX evolution. Therefore, a 1:1 mixture of *Taq/Pwo* was prepared and another pilot PCR performed on 5xA/0xP XL-digest.

It was theorized that since *Pwo* cannot amplify DNA sequences with cross-link fragments, it would essentially do nothing during the first cycle of PCR. Conversely, *Taq* might replicate these sequences during the first few cycles of PCR and thereby provide unperturbed templates for amplication by *Pwo* in subsequent

PCR cycles. This approach resulted in substantially improved yields with far fewer aborted fragments as shown in Figure R.1.5 on below.

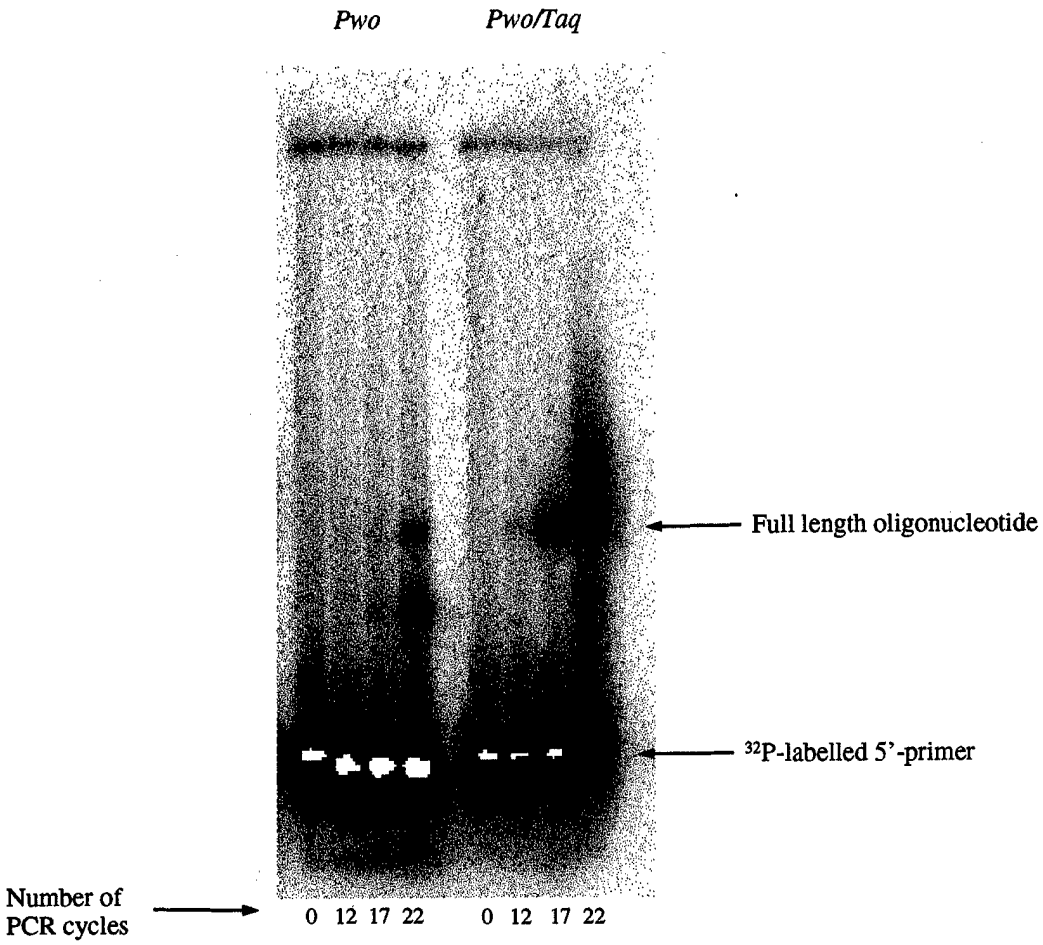


Figure R.1.5. PAGE of pilot PCR amplification of 5xA/0xP XL-dig using *Pwo/Taq* in comparison to *Pwo* amplification alone.

With the apparent success of PCR amplification of cross-link/digest aptamers, it was now possible to evaluate the evolution of a PhotoSELEX experiment. As mentioned in the Introduction, one criterion of assessing PhotoSELEX evolution is cross-link yield. Cross-link yield can be determined from a phosphor-image of the PAGE partition of a cross-link reaction by measuring the activity present in the

photocross-link band and dividing that by the total activity present in that lane. After one round of PhotoSELEX, a cross-link signal was observed for the *Pwo/Taq* amplified library which supported the premise that *Taq* could indeed amplify past the cross-link site. Unfortunately, as illustrated in Figure R.1.6 below, an overall reduction in cross-link yield was detected.

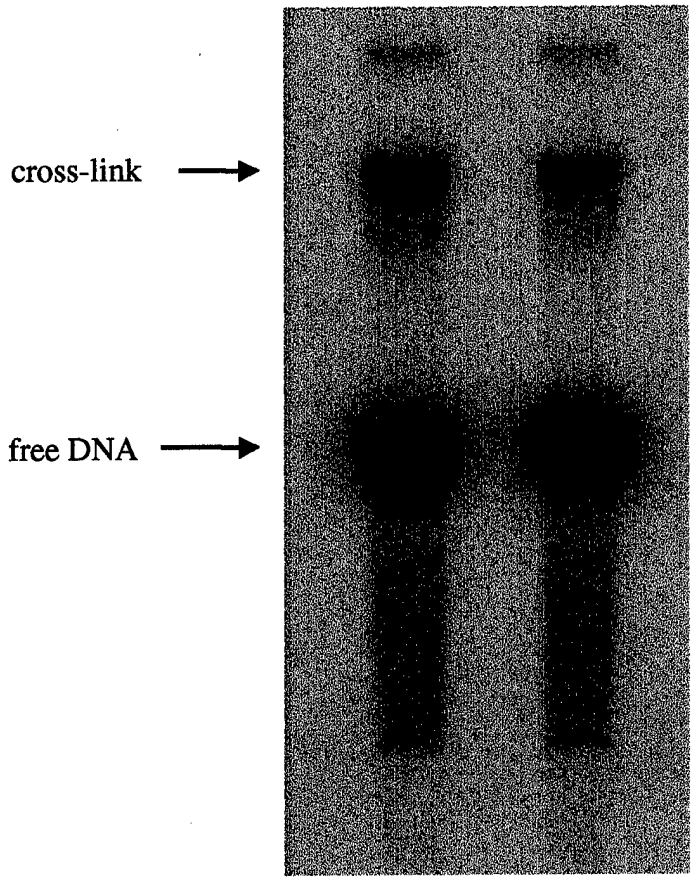


Figure R.1.6. Evolution of cross-link yield after one round of PhotoSELEX when selecting the apparent cross-link band.

Obviously, this immediately called into question the feasibility of the

PhotoSELEX concept. One possible explanation for the reduced cross-link yield was an imperfect (albeit improved over *Pwo*) amplification through the cross-link fragment by *Taq*. Even a fairly small mutation rate at the cross-link site could result in a perceptible reduction in cross-link yield from one round to the next and ultimate failure for PhotoSELEX. Another explanation, however, was that the original library possessed substantial nonspecific cross-link; once the background had been

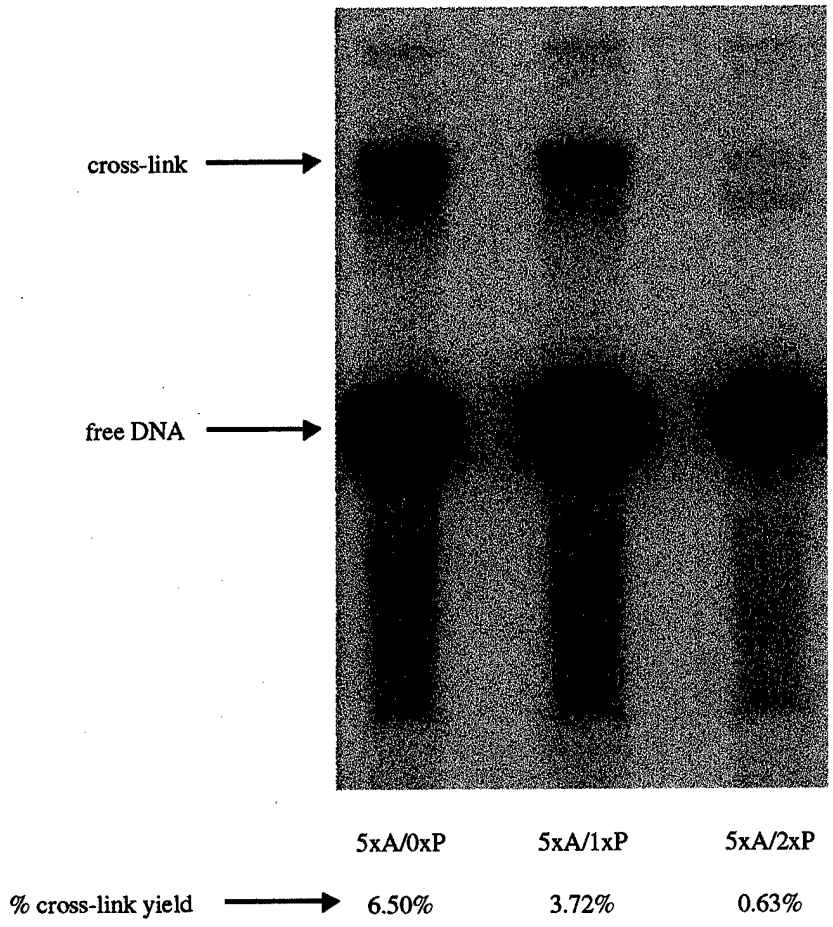


Figure R.1.7. Evolution of cross-link yield after two rounds of PhotoSELEX when selecting the apparent cross-link band.

eliminated, PhotoSELEX should show improving cross-link yields. With this notion in mind, another round of PhotoSELEX was completed. However, as Figure R.1.7 on the previous page illustrates, the cross-link yield had decreased again, and it was becoming difficult to discern the presence of a cross-link signal. Background confounding of the cross-linking aptamers did not appear to be the problem. Yet, the question of background/nonspecific cross-linking suggested a third possible explanation.

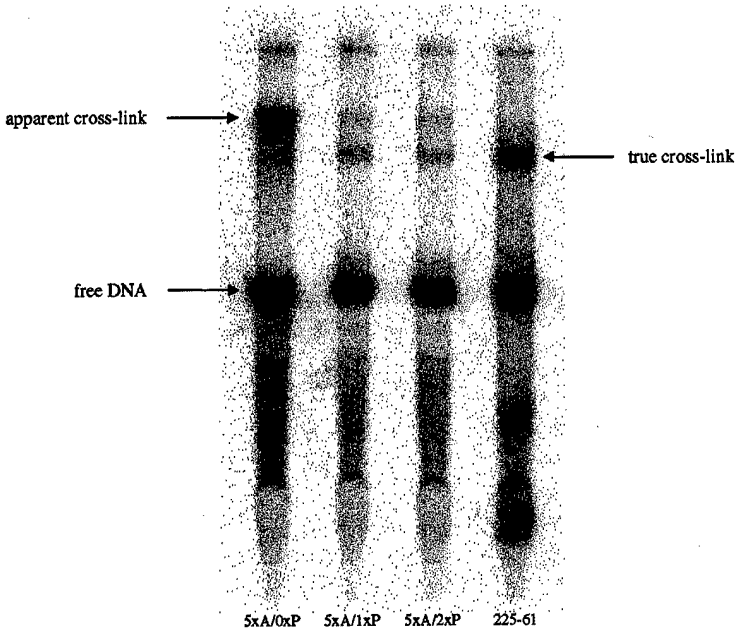


Figure R.1.8. Determination of the true on molecule DNA/one molecule bFGF cross-link band (50 nM DNA/50 nM bFGF/250 pulses 308 nm light).

As has been mentioned several times, partitioning is a key component of the rate of SELEX convergence. The more effective the partitioning mechanism is at separating the desired sequences from the undesired sequences, the more rapid will be the convergence. However, partitioning will also break a SELEX if it does not return

the desired sequences. Certainly one way to effect this in PhotoSELEX is to excise gel bands which do not contain legitimate cross-linking sequences. To verify that the correct bands were being excised during PAGE partitioning, the post-affinity SELEX modifed aptamer, 225-61, was cross-linked with bFGF and PAGE partitioned along with the bFGF/cross-linked libraries corresponding to rounds 0,1 and 2 of PhotoSELEX. As can be clearly seen in Figure R.1.8 on the previous page, the region corresponding to true one molecule of DNA/one molecule of bFGF cross-linking had been discarded during partitioning in the initial PhotoSELEX rounds in favor of a substantial, but primarily background cross-link band.

Finally, a PhotoSELEX round was completed for the bFGF target using the fifth library of an affinity SELEX experiment for bFGF in which T had been replaced by BrdU. This ssDNA library employing a 30-mer random region was photocross-linked with bFGF at 50 nM DNA/10 nM bFGF/250 pulses of 308 nm light. After subsequent PAGE partitioning and excision of a region corresponding to the migration position of 225-61/bFGF cross-linked adduct, digestion of bFGF with Proteinase K, and amplification of the cross-linking sequences with *Taq-Pwo*/PCR, the library was photocross-linked again with bFGF under identical conditions. As can be seen from Figure R.1.9 on the following page, the desired cross-link signal improved substantially after one PhotoSELEX round.

It now seemed possible to continue PhotoSELEX rounds in precisely this manner until convergence was achieved. The degree of evolution was monitored by measuring the cross-link yield by phosphor-imaging for each round, and it was anticipated that the evolved DNA pool would show substantial cross-link yield with bFGF. However, after only three rounds of PhotoSELEX, further photoselection

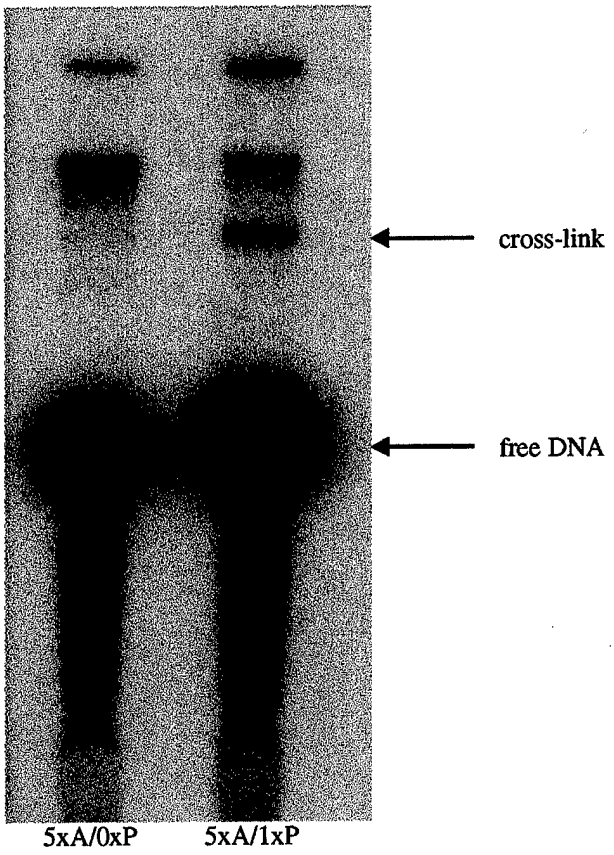


Figure R.1.9. One round of successful PhotoSELEX with partitioning based on true one molecule DNA/one molecule bFGF cross-link (50 nM DNA/10 nM bFGF/250 pulses 308 nm light).

employing a 50 nM/10 nM oligonucleotide to protein ratio and 250 pulses of 308 nm light irradiation did not result in further evolution. Since Irvine et al. (1991) argue that the rate of convergence in affinity SELEX will be substantially slowed without increasing oligonucleotide binding competition, similar reasoning would suggest that PhotoSELEX would not continue to evolve without increasing the binding and/or photochemical competition. In an effort to evaluate the relative contribution of these two pressures to PhotoSELEX convergence, they were made independently more stringent by branching the PhotoSELEX experiment. One branch increased the

affinity pressure by either reducing protein concentration while keeping DNA concentration constant or by keeping protein concentration constant while increasing DNA concentration--all while keeping the number of photons delivered to the cross-linking reaction constant. The other branch sought to increase photochemical pressure by reducing the number of photons delivered while keeping DNA and protein concentration constant. This PhotoSELEX scheme is shown in Figure R.1.10 below.

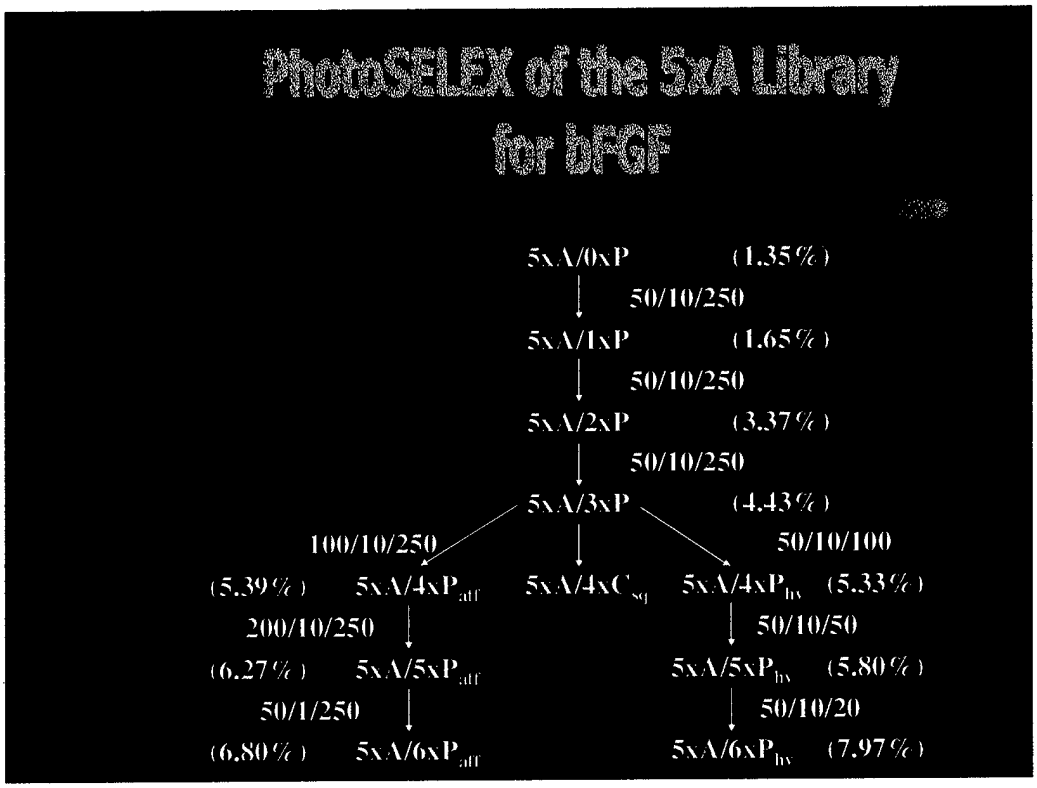
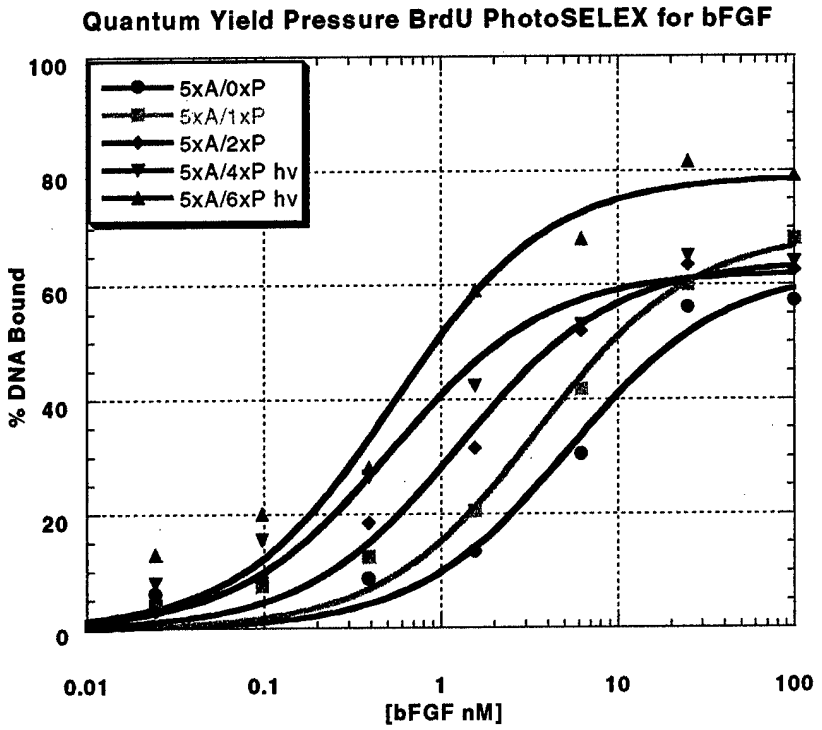


Figure R.1.10. PhotoSELEX scheme for the 5xA library. (Library % cross-link measured at 50 nM DNA/10 nM bFGF/250 pulses 308 nm light. PhotoSELEX condition indicated as x/y/z: x = [DNA], y = [bFGF], z = pulses 308 nm light.)

As mentioned earlier, an important criterion by which PhotoSELEX evolution

can be monitored is the phosphor-image measurement of the bulk photocross-link yield for the various libraries. Figure R.1.10 indicates the photocross-link yield for the six libraries resulting from both the affinity and photochemical pressure experiments (reported in parentheses). The photocross-link reaction used to report these yields was performed at the standard cross-link condition (50 nM DNA/10 nM bFGF/250 pulses 308 nm light). It should be noted that the final library resulting from quantum yield pressure displays somewhat greater photocross-linking ability than that evolved from affinity pressure. While this makes sense intuitively (increased photochemical pressure should yield the most fascile photocross-linking aptamers), the difference is not large and may be a result of experimental error. Nevertheless, on the basis of the cross-link criterion, the PhotoSELEX protocol apparently succeeded in evolving cross-linking sequences to bFGF when either affinity *or* photochemical pressure was applied. However, it was unknown whether or not PhotoSELEX had simultaneously succeeded in yielding high-affinity binders for bFGF.

A first assessment of this capability was provided by measuring the K_ds of the various libraries which PhotoSELEX had evolved. As can be seen in Figures R.1.11 and R.1.12 on the following two pages, both the affinity branch *and* the quantum yield branch evolved high-affinity binders. As anticipated, the affinity branch resulted in a final library with better affinity for bFGF than the quantum yield branch; this result is consistent with the observation that the quantum yield branch resulted in a final library with the best cross-linking ability. At this point, it seemed PhotoSELEX had done everything that its originators hoped it would: the evolution of high-affinity binding and high quantum yield photocross-linking aptamers for a given

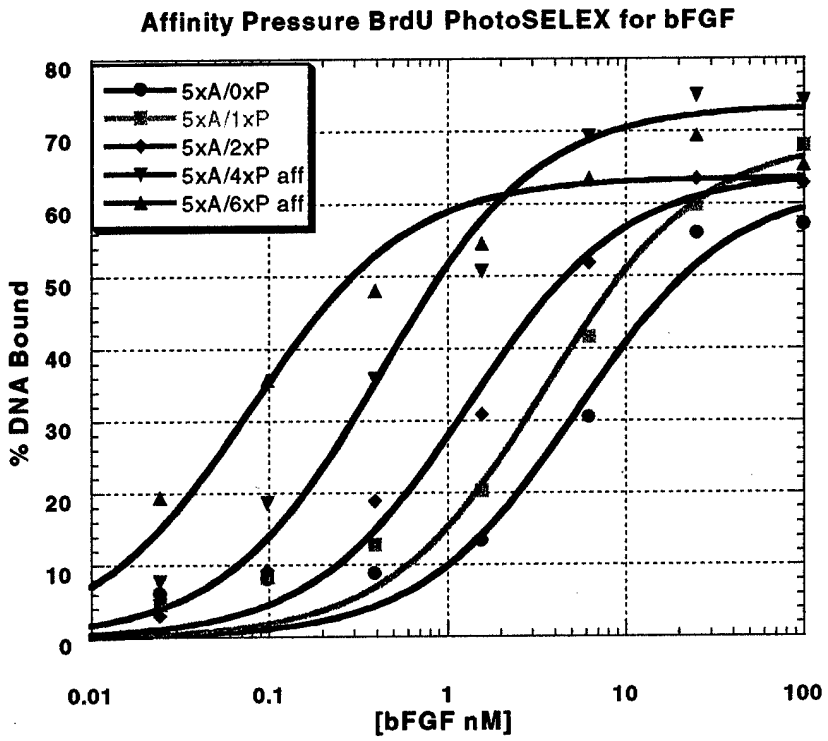


Library	5xA/0xP	5xA/1xP	5xA/2xP	5xA/4xP hv	5xA/6xP hv
K_d	5.3 nM	3.5 nM	1.3 nM	0.49 nM	0.50 nM

Figure R.1.11. Evolution of the quantum yield pressure BrdU PhotoSELEX for bFGF as measured by library affinity.

target. Moreover, these results indicate PhotoSELEX had exceeded expectation with respect to identification of high affinity-binding aptamers (a library K_d of 0.08 nM is so low as to virtually guarantee the presence of very-high-affinity binding aptamers). The apparent success of PhotoSELEX to evolve such high-affinity binders upon stringent affinity selection prompted a more thorough evaluation of PhotoSELEX to evolve high-quantum yield photocross-linkers upon stringent photochemical selection.

Since the standard photocross-linking reaction used to evaluate PhotoSELEX



Library	5xA/0xP	5xA/1xP	5xA/2xP	5xA/4xP aff	5xA/6xP aff
K_d	5.3 nM	3.5 nM	1.3 nM	0.42 nM	0.08 nM

Figure R.1.12. Evolution of the affinity pressure BrdU PhotoSELEX for bFGF as measured by library affinity.

evolution employs a five-fold DNA excess to protein, it was assumed that performing a 1:1 DNA/protein reaction should result in five times the cross-link yield for the same number of pulses. To test this argument, a time course photocross-linking reaction was completed with 50 nM DNA (5xA/6xP hv)/50 nM bFGF. As shown in Figure R.1.13 on the following page, while substantial cross-link yield was observed, it was not five times that observed for the standard reaction. In fact, even irradiation for 2000 pulses did not yield quite five times that seen for 5xA/6xP hv at 250 pulses

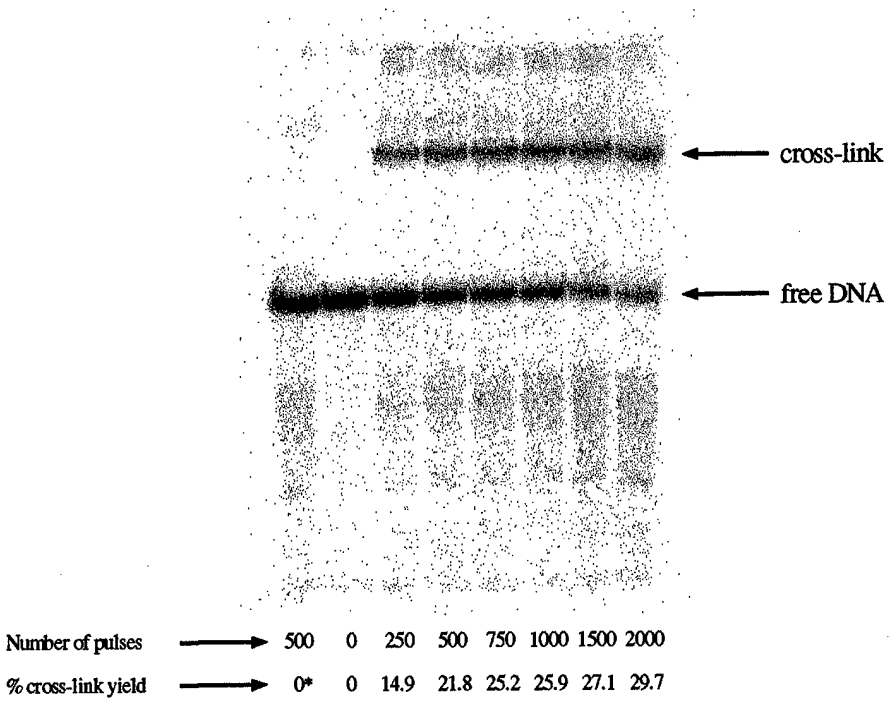


Figure R.1.13. Time course of 5xA/6xP library cross-linked with bFGF at 50 nM DNA/50 nM bFGF (* indicates no bFGF).

with 5-fold DNA excess. The inability to extrapolate yields from five-fold DNA excess to a 1:1 ratio may be explained by photochemical damage to the oligonucleotide. In the case of five-fold DNA excess, a damaged copy of a sequence presumably has four other copies available to still generate the cross-link. However, this explanation does not account for the observation that increasing the number of pulses for the 1:1 ratio reaction increases the cross-link yield. A more probable explanation is that DNA excess forces competition between the members of the library with subsequent binding and cross-linking of the very best aptamers. When the stringency is relaxed, more aptamers which cross-link with lower quantum yield bind bFGF molecules with subsequent lower cross-link yield than that predicted by linear extrapolation. If this observation is indeed an artifact of the library, the linear

extrapolation should hold once individual aptamers are identified. Nevertheless, these cross-link yields are excellent and rival commonly reported yields employing the BrdU chromophore (see Introduction). Furthermore, these conspicuously positive results suggested that the true potential of PhotoSELEX had not yet been determined.

As mentioned at the outset, it was initially hypothesized that PhotoSELEX would not be capable of evolving high-affinity binding/high quantum yield cross-linking aptamers from a completely randomized pool. The basis of this hypothesis is that the conditions required for photocross-linking are so specific as to preclude adequate sampling of the sequence space in a completely random oligonucleotide library (regardless of the fact that the protein concentration can made greater than the DNA concentration). However, in view of the rapid convergence facilitated by PhotoSELEX when starting from the 5xA library, it certainly seemed that using PhotoSELEX to evolve desired aptamers without the benefit of prior affinity selection or bias was an idea worth exploring.

The next PhotoSELEX experiment was performed with a completely random DNA library in which T had been replaced with BrdU. As with the 5xA-PhotoSELEX experiment, deciding on the DNA/protein ratio would be crucial to the ultimate success of *de novo* PhotoSELEX (perhaps even more so in this case). Again, the formula for near optimal protein concentration established by Irvine et al., (1991) was used for this determination. While it was originally assumed that the partition efficiency as expressed by the ratio (BG/CP) was 10-15%, it now seemed this may have been somewhat overestimated (PAGE partition probably yields (BG/CP) of no greater than 5-6%). Therefore, starting with a library K_d of 200 nM, taking as a standard DNA concentration 50 nM, and assuming a partition efficiency of 5-6%,

T_{no}= 25nM. As a result, the first round of 0xA PhotoSELEX was completed at 50 nM DNA/25 nM bFGF/250 pulses 308 nm light with the goal in this PhotoSELEX experiment to evolve high-affinity binding/high quantum yield cross-linking aptamers as rapidly as possible.

The first PhotoSELEX experiment showed that both affinity pressure and photochemical pressure were capable of evolving desired sequences relative to the pressure applied. The first experiment also showed continually increasing stringency facilitated a reasonable rate of evolution. Therefore, it was decided to couple the affinity and photochemical selective pressures for the 0xA PhotoSELEX experiment. As shown in Figure R.1.14 on below, this was done by decreasing the protein

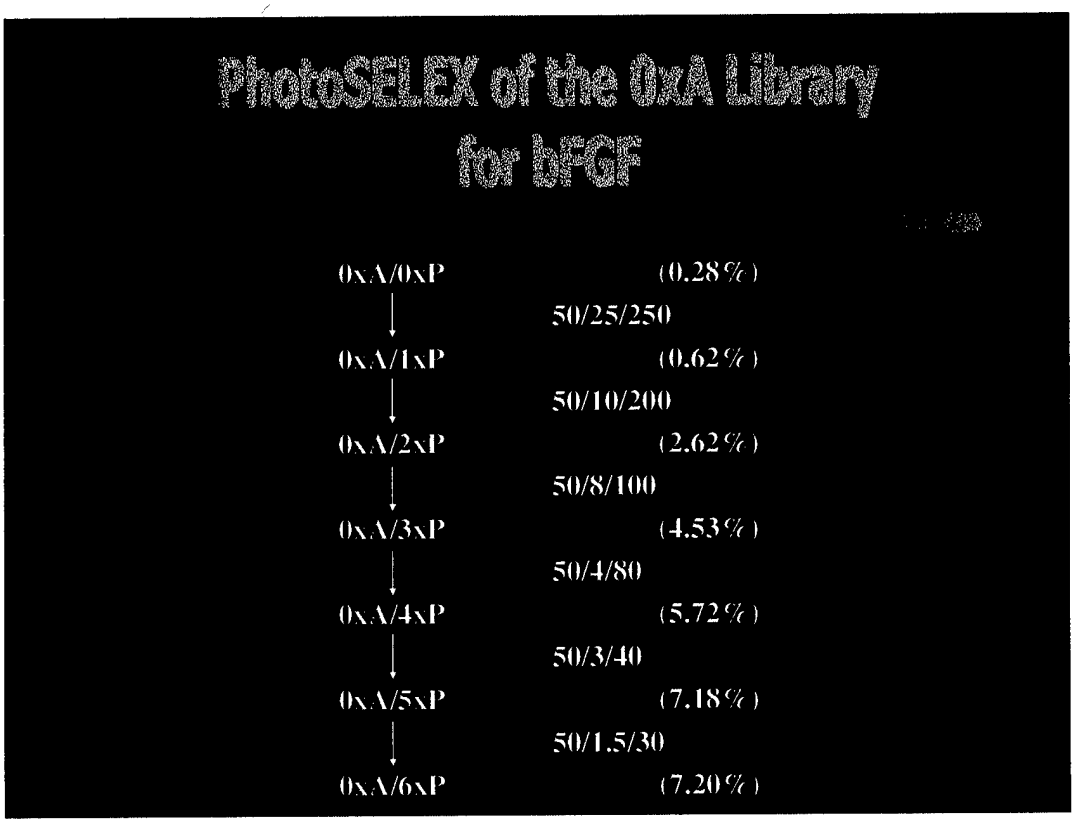
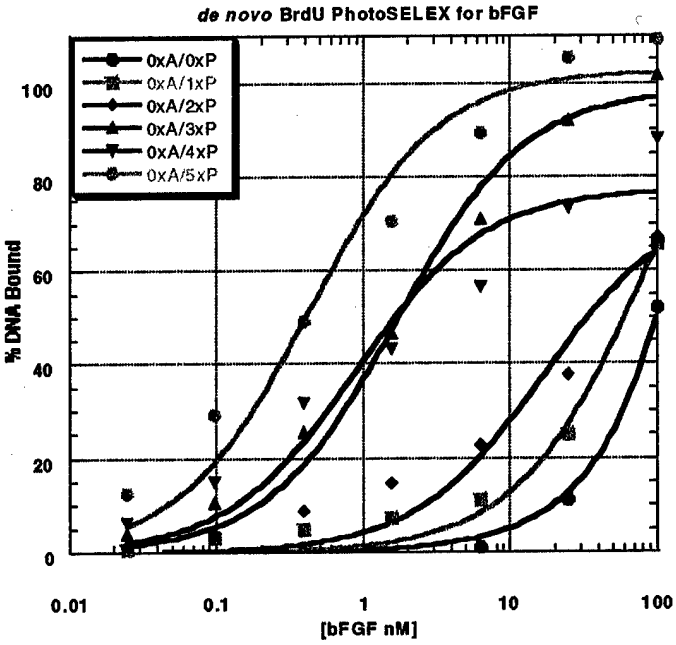


Figure R.1.14. PhotoSELEX scheme for the 0xA library. (Library % cross-link yield and PhotoSELEX conditions reported here as in Figure R.1.10.)

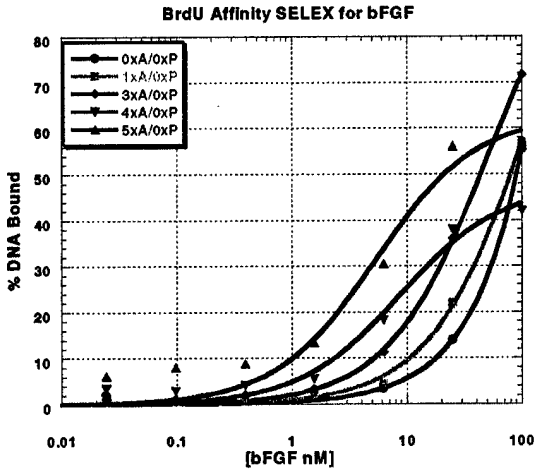
concentration and photons delivered in every round. However, it should be noted that both pressures were not increased to the same extent each time. Rather, in order to minimize the potential for breakage of the PhotoSELEX protocol, affinity or photochemical stringency was emphasized in alternating rounds. For instance, a large increase in affinity pressure was applied in proceeding from round 1 (0xA/1xP) to round 2 (0xA/2xP) whereas a relatively small increase in photochemical pressure was applied. Conversely, a large increase in photochemical pressure was applied in proceeding from round 2 to round 3 (0xA/2xP to 0xA/3xP) while simultaneously applying a fairly small increase in affinity pressure. As Figure R.1.14 on the previous page shows, following this protocol resulted in a substantial increase in library cross-link yield with the evolved library (0xA/6xP) exhibiting very similar cross-link yield to 5xA/6xP hν/aff under the same conditions (50 nM DNA/10 nM bFGF/250 pulses 308 nm light).

A measurement of the 0xA PhotoSELEX library K_ds revealed an increase in binding affinity for bFGF similar to that observed for the 5xA PhotoSELEX experiment as seen Figure R.1.15 on the following page. While the K_d of 0xA/6xP is not quite as good as 5xA/6xP aff, it nevertheless indicates very good binding affinity. What is very impressive, however, is the rate of evolution as indicated by the library K_ds. Starting from the same completely random oligonucleotide library, after five rounds, PhotoSELEX has evolved a library with an affinity an order of magnitude better than that evolved after five rounds of affinity SELEX. Figure R.1.1 is duplicated here to facilitate comparison of *de novo* PhotoSELEX and affinity SELEX rates of evolution.



Library	0xA/0xP	0xA/1xP	0xA/2xP	0xA/3xP	0xA/4xP	0xA/5xP
K_d	200 nM	84 nM	17 nM	1.7 nM	0.89 nM	0.42 nM

Figure R.1.15. Evolution of *de novo* BrdU PhotoSELEX for the bFGF target as measured by library affinity.



Library	0xA/0xP	1xA/0xP	3xA/0xP	4xA/0xP	5xA/0xP
K_d	200 nM	124 nM	51 nM	8.6 nM	5.3 nM

Figure R.1.1. Evolution of BrdU affinity SELEX for the bFGF target as measured by library affinity.

All of these results support the following conclusions about PhotoSELEX:

1. PhotoSELEX can be accomplished against a target for which no advantage derived from biased randomization or prior affinity selection is avaliable.

2. PhotoSELEX appears to be capable of identifying the best photocross-linking aptamers among all aptamers which cross-link a given target.

3. PhotoSELEX can truly select for affinity while selecting for photocross-linking ability.

4. PhotoSELEX can be performed in a systematic, reproducible manner without *in situ* modifications.

As such, all four questions posed in the Introduction regarding PhotoSELEX have been answered in the affirmative. One may reasonably conclude that PhotoSELEX can augment if not supplant affinity SELEX as a combinatorial chemistry technique for the high-throughput identification of diagnostic and therapeutic agents. This seems particularly plausible when presented with the result that PhotoSELEX can evolve high-affinity aptamers more rapidly from a completely unevolved library than can affinity SELEX.

CHAPTER IV: CLONING AND SEQUENCING OF THE PHOTOSELEX LIBRARIES

While the binding affinity and photocross-linking yields of the libraries evolved by PhotoSELEX strongly indicated excellent evolution of desired aptamers, the ultimate objective of any SELEX experiment is the identification of the very best performing individual sequences with respect to the chosen target. This identification begins with cloning and sequencing of the evolved libraries. The oligonucleotide sequences identified by cloning and sequencing of the three evolved PhotoSELEX libraries are shown on the following three pages. Within each library, the various sequences are organized into families that share substantial sequence homology. Each sequence is shown with the entirety of its random region in upper case letters and base-specific coloring to aid visualization of homology. Bases in the fixed region which are thought to be important in base pairing are appended to the random region and shown in lower case lettering. Finally, base pairing identified by either computer folding or analysis of covariance, as described later, is identified by underlining.

As can be seen immediately, the degree of sequence convergence and homology is outstanding. In fact, among all three libraries, there appear to be just two primary families. Family I is composed of one sequence which does not seem to share any obvious similarity with members of Family II. Since the Family I sequence appears multiple times in each library, a discussion regarding its potential import to the PhotoSELEX experiments and its possible secondary structure is certainly warranted.

The most striking feature of the Family I sequence (from here on referred to as 06.50--the 50th sequence identified from the 0xA/6xP library), is that it appears identically, multiple times in each evolved library. It is difficult to comprehend why no

TABLE R.2.1
Sequences of BrdU aptamers for bFGF from the 5xA/6xP hν library
(T = BrdU)
gggaggacgatgcgg | 30N | cagacgacgagcggga

Family I		
56hν.6	tgcggTGACGTAAGAGTGTAATCGATGCAGCCTGGcagacgacgagcg	
56hν.49	tgcggTGACGTAAGAGTGTAATCGATGCAGCCTGGcagacgacgagcg	
56hν.30	tgcggTGACGTAAGAGTGTAATCGATGCAGCCTGGcagacgacgagcg	
56hν.m1	tgcggTGACGTAAGAGTGTAATCGATGCAGCCTGGcagacgacgagcg	
56hν.m7	tgcggTGACGTAAGAGTGTAATCGATGCAGCCTGGcagacgacgagcg	
56hν.m12	tgcggTGACGTAAGAGTGTAATCGATGCAGCCTGGcagacgacgagcg	
56hν.m15	tgcggTGACGTAAGAGTGTAATCGATGCAGCCTGGcagacgacgagcg	
56hν.m22	tgcggTGACGTAAGAGTGTAATCGATGCAGCCTGGcagacgacgagcg	
Family II		
56hν.8	tgcggACCACAGCGCTGATGG	TTCATACTATTCCGca
56hν.14	tgcggACCACAGCGCTGATGG	TTCATACTATTCCGca
56hν.15	tgcggACCACAGCGCTGATGG	TTCATACTATTCCGca
56hν.m20	tgcggCCAAGGTGACATCCGGG	TTCATAGTATCCGca
56hν.29	GCCGAAGTGACACGAGG	TTCATAGTATGCC
56hν.m11	tgcggCGCAAGCAAACGCTGAG	TTCATAGTATCCGca
56hν.7	tgcggACCACAGCGATGATGG	TTCATATTATTCCAca
56hν.20	tgcggACCACAGCGATGATGG	TTCATATTATTCCAca
56hν.m4	tgcggGCGAAGTCATTCACGAG	TTCATCATATCCCcaga
56hν.2	GCAAAGTTCTACGAG	TTCATCTTATCCCAA
56hν.m23	GCAAAGTTCTACGAG	TTCATCTTATCCCAA
56hν.41	GCAAAGTTCTACGAG	TTCATCTTATCCCAA
56hν.11	tgcggGCCAAGGGTCTCACGG	TTCATTCTATCCCAca
56hν.50	ggGGCAAGCCAACGCGTG	TTCATTGTATCCCC
56hν.46	tgcggCAAAGACATACTGGAGG	TTCACTTTATCCGca
56hν.32	ACCAAAGCCAGCGGG	TTACATACTATCCTG
56hν.26	GTACGAAGGCGACCCGAG	TTAATCTTACCC
56hν.33	ACATGAAGCTACTGCGAC	TTAAATCTTGGC
56hν.m21	GAACAAAGGCGCCCGTGTA	TTTTCTTTCCC
56hν.38	GAACAAAGGCGCCCGTGTA	TTTTCTTTCCC
56hν.4	CGAGCGGAGACG	TTTTATTATCCA
56hν.17	CGAGGTAGGCCAG	TTTTATTATCA

TABLE R.2.2
Sequences of BrdU aptamers for bFGF from the 5xA/6xP aff library
(T = BrdU)

gggaggacgatgcgg |30N| cagacgacgagcggga

Family I		
56af.6	tgcggTGACGTAAGAGTGTAATCGATGCAGCCTGGcagacgacgagcg	
56af.13	tgcggTGACGTAAGAGTGTAATCGATGCAGCCTGGcagacgacgagcg	
56af.15	tgcggTGACGTAAGAGTGTAATCGATGCAGCCTGGcagacgacgagcg	
56af.32	tgcggTGACGTAAGAGTGTAATCGATGCAGCCTGGcagacgacgagcg	
56af.36	tgcggTGACGTAAGAGTGTAATCGATGCAGCCTGGcagacgacgagcg	
56af.38	tgcggTGACGTAAGAGTGTAATCGATGCAGCCTGGcagacgacgagcg	
Family II		
56af.19	TgcggACCACAGCGCTGATGG	TTCATACTATTCCGca
56af.20	TgcggACCACAGCGCTGATGG	TTCATACTATTCCGca
56af.45	TgcggACCACAGCGCTGATGG	TTCATACTATTCCGca
56af.47	TgcggACCACAGCGCTGATGG	TTCATACTATTCCGca
56af.37	CAGGCTACCACAGTACCGAG	TTCATACTAT
56af.40	CAACCACTGGGGCTTG	TTCATAGTATCCGC
56af.25	GCCGAATCATTGAGAG	TTCATAGTATGCCC
56af.22	CGAAGGGACAACCCTAC	TTCATAATATCCG
56af.3	GCCGAAGTTCTAACGCG	TTCATCCTATGCC
56af.21	GCCGAAGTTCTAACGCG	TTCATCCTATGCC
56af.16	GTCGAAGCCATGCAAG	TTCATTGTATACCC
56af.41	GTCGAAGCCATGCAAG	TTCATTGTATACCC
56af.23	CGAAGGTCACCGAG	TTCTATGCTATCCGCA
56af.30	CGAAGGTCACCGAG	TTCTATGCTATCCGCA
56af.4	CGAAGGTCACCGAG	TTCTGTGCTATCCGCA
56af.17	CCGGAGGTCTCCAAG	TTCATTACTATGCCA
56af.5	GACGAAGGCGTCCGAG	TTCAATCTTCCCAG
56af.27	GACGAAGACGTGCGAG	TTGAATCTTCCCAG
56af.29	GACGAAGGCGTCCGAG	TTGAATCTTCCCAG
56af.1	ggCCACGAAGGCGACCCGAG	TTTCATCTTGGCc
56af.46	ggCCACGAAGGCGACCCGAG	TTTCATCTTGGCc
56af.12	GCCGAAGTTCTAACGCG	TTTATCCTATGCC
56af.2	CGAAAGGCAACCCGAG	TTTATAGTATCCAG
56af.26	GCGAAGGCACACCGAG	TTTATAGTATCCCA

TABLE R.2.3
Sequences of BrdU aptamers for bFGF from the 0xA/6xP library
(T = BrdU)

gggaggacgatgcgg | 30N | cagacgacgagcggga

Family I	
06.24	tgcggTGACGTAAGAGTGTAATCGATGCAGCCTGGcagacgacgagcg
06.30	tgcggTGACGTAAGAGTGTAATCGATGCAGCCTGGcagacgacgagcg
06.32	tgcggTGACGTAAGAGTGTAATCGATGCAGCCTGGcagacgacgagcg
06.36	tgcggTGACGTAAGAGTGTAATCGATGCAGCCTGGcagacgacgagcg
06.45a	tgcggTGACGTAAGAGTGTAATCGATGCAGCCTGGcagacgacgagcg
06.49	tgcggTGACGTAAGAGTGTAATCGATGCAGCCTGGcagacgacgagcg
06.50	tgcggTGACGTAAGAGTGTAATCGATGCAGCCTGGcagacgacgagcg
06.58	tgcggTGACGTAAGAGTGTAATCGATGCAGCCTGGcagacgacgagcg
Family II	
06.1	tgcggACCACAGCGCTGATGG TTCATACTATTCCAca
06.43	tgcggACCACAGCGCTGATGG TTCATACTATTCCAca
06.3	tgcggACCAAAGGCAATCCGGG TTCATACTATTCCcaga
06.53	tgcggACCAAAGGCAATCCGGG TTCATACTATTCCcaga
06.15	tgcggGCGAAGGCACACCGAG TTCATAGTATCCCAca
06.19	tgcggGCGAAGGCACACCGAG TTCATAGTATCCCAca
06.41	tgcggGCGAAGGCACACCGAG TTCATAGTATCCCAca
06.10	tgcggCGCAAGCAAACGCTGAG TTCATAGTATCCGca
06.44	tgcggCGCAAGCAAACGCTGAG TTCATAGTATCCGca
06.55	tgcggGGCAAGCGAACGCTGAG TTCATAGTATCCGca
06.12	gTGACGAAGGCTACCGAG TTCATATTATTCAc
06.17	tgcggCGAAGGCTACCTCCAAG TTCATATTATCCGca
06.20	GGCCATGTCTCATAG TTCATATTATACCTG
06.59	GGCCATGTCTCATAG TTCATATTATACCTG
06.35	ggGCCGAAGTTCTAACGCG TTCATCCTATGCCc
06.52	ggGCCGAAGTTCTAACGCG TTCATCCTATGCCc
06.57	ggGCCGAAGTTCTAACGCG TTCATCCTATGCCc
06.5	tgcggCGAAGGTCACCGAG TTCTATGCTATCCGCA
06.11	tgcggCGAAGGTCACCGAG TTCTATGCTATCCGCA
06.18	tgcggCGAAGGTCACCGAG TTCTATGCTATCCGCA
06.7	ggCCACGAAGGCGACCCGAG TTTCATCTTGGCc
06.22	ggCCACGAAGGCGACCCGAG TTTCATCTTGGCc
06.31	ggCCACGAAGGCGACCCGAG TTTCATCTTGGCc
06.60	ggCCACGAAGGCGACCCGAG TTTCATCTTGGCc
06.28	ggGACGAAGGCGCCCGAG TTGAATCTTCCCAG
06.29	ggGACGAAGGCGCCCGAG TTGAATCTTCCCAG
06.38	ggGACGAAGGCGCCCGAG TTGAATCTTCCCAG

sequences which vary even slightly from this sequence did not also evolve in at least one other library; only two explanations seem possible. The less optimistic (but seemingly more plausible) of these explanations is that some aspect or process within the PhotoSELEX protocol continually introduced 06.50 as a contaminant. For instance, if a number of copies of this sequence had somehow found their way into the dNTP solution used for PCR amplification during each round of PhotoSELEX, this sequence would contaminate each amplified library. Although conceiving of a source of this previously unknown sequence is difficult, the absolute lack of any variance would seem to implicate it as an impurity. The more optimistic of these explanations (but decidedly less plausible) is that 06.50 is simply the highest quantum yield cross-linking/highest affinity binding aptamer available from all those present in any starting library. Somehow this one sequence so completely embodies the best combination of cross-linking and binding that even one modified base position significantly reduces these attributes. Obviously, screening the sequences will indicate whether 06.50 is a legitimately-evolved aptamer or whether it is a contaminant. However, another question regarding 06.50 (and several of the Family II sequences) is why it appears in all three libraries.

Typically, two SELEX experiments performed for the same target will usually evolve similar but not identical aptamers. When identical aptamer sequences do evolve, it is usually considered a consequence of inter-library contamination. In inter-library contamination, sequences evolved for a target make their way into a subsequent experiment for the same target as gel plate residues, through shared reagents, on common surfaces, and even via air-borne transport. Once a winning-sequence for a

given target with the same 5' and 3' fixed regions infiltrates a SELEX experiment, it will evolve just as if it had been present from the outset. Therefore, the fact that 06.50 appears in all three libraries calls into question the claim that PhotoSELEX is capable of evolving high-quantum yield cross-linking/high-affinity binding aptamers without the benefit of prior affinity selection. Fortunately, in the event that 06.50 proves to be a high quantum yield cross-linking/high affinity binding aptamer for bFGF, there is an alternative explanation for its presence in all three libraries that is consistent with PhotoSELEX as a viable *de novo* combinatorial chemistry protocol.

The presence of 06.50 in all three libraries can be explained by PhotoSELEX's exquisite ability to identify the perfect solution to the applied photochemical/affinity pressures. However, this can only be true if at least two copies of 06.50 were initially synthesized during the preparation of the completely random DNA library. Since the first PhotoSELEX libraries (5xA/6xP hν and 5xA/6xP aff) were evolved from a DNA library which itself had been through five rounds of affinity SELEX, at least one copy of 06.50 must have been present in the random synthetic library from which it was evolved. However, the final PhotoSELEX library (0xA/6xP) was evolved from this same original synthetic library. Therefore, the only way 06.50 could have been legitimately evolved in the 0xA/6xP library is if at least one other copy of 06.50 was still present in the random synthetic library after the fraction required for the affinity SELEX experiment had been removed.

This at first may not seem a likely occurrence. For instance, as described in the Introduction, there are 1.153×10^{18} possible unique sequences for a library incorporating a 30-mer random region (i.e. 4^{30}), and the chance of synthesizing the

same one at least twice appears exceptionally improbable. However, the generation of the random, synthetic oligonucleotide library was completed on a 1 µmol scale. Assuming 100% yield, a 1 µmol synthesis will theoretically generate 6.022×10^{17} sequences, or, put another way, 6.022×10^{17} chances to generate *at least two* matching sequences. If it is assumed that every possible sequence has an equal likelihood of being synthesized $(1/4^{30})$, the problem can be viewed essentially as a binomial probability distribution. If the synthesis of a particular sequence like 06.50 is considered a success (S) and any other sequence is considered a failure, (F) the probability of generating *at least two matching sequences* is given by:

$$1 - P(X \leq 1)$$

$$where \quad P(X \leq 1) = \sum_{x=0}^{1} b\left(x; 6.022 \times 10^{17}, 1/4^{30}\right)$$

$$b\left(x; 6.022 \times 10^{17}, 1/4^{30}\right) = \binom{6.022 \times 10^{17}}{x} \left(1/4^{30}\right)^{x} \left(1 - 1/4^{30}\right)^{6.022 \times 10^{17} - x}$$

$$and \binom{6.022 \times 10^{17}}{x} = \frac{6.022 \times 10^{17}!}{x!\left(6.022 \times 10^{17} - x\right)!}$$

This computation is relatively straightforward except for the term $\left(1 - 1/4^{30}\right)^{6.022 \times 10^{17}}$ which was computed using Mathematica ver. 3.0 software (Wolfram Research Inc.). From this calculation, the probability of generating at least two identical sequences is 69%. This idealized computation illustrates that it is *possible* to evolve 06.50 and other sequences which appear in all three PhotoSELEX-evolved libraries by legitimate selection. Moreover, Osborne and Ellington (1997) argue that synthetic limitations reduce the number of different sequences that can actually be generated by solid phase oligonucleotide synthesis. As such, there may be many more

copies of certain sequences than the above calculation would indicate; this increases the likelihood that these multiple-appearing sequences did not result from inter-library contamination. Finally and most convincingly, Collins and coworkers have completed *de novo* PhotoSELEX experiments on novel protein targets with successful convergence (personal communication, 1998). It can now be asserted with some confidence that PhotoSELEX can identify the very highest quantum yield photocross-linking/highest affinity binding aptamers from a completely randomized oligonucleotide library. While this settles the argument about the evolution of identical sequences, it does not answer the question as to whether 06.50 is an impurity or a legitimate aptamer. While screening the aptamers will unequivocally resolve the issue, examination of 06.50's secondary structure may help predict the outcome.

Elucidation of an aptamer's secondary structure is aided by regions of covariance. Covariance is defined as conserved sequence variation among aptamers of an evolved family. For example, the first three sequences of a family might possess a T at a certain position with an A at some distant locus. If the fourth sequence includes a G in place of T and C in place of A, then this would indicate an important covariant region of base pairing essential for secondary structure. Unfortunately, such an analysis for 06.50 is not possible. Therefore, a first guess as to secondary structure for this sequence is provided by computer folding programs which attempt to minimize Gibbs Free Energy by maximizing base pairing for a given sequence. In most programs such as MulFold used here, ssDNA is viewed equivalent to RNA (Zucker 1989; Jaeger et al., 1989a; Jaeger et al., 1989b). Figure R.2.1 on the following page shows three reasonable structures generated by MulFold for the 06.50 sequence. Note

that all three structures incorporate at least one chromophore at a protein-accessible hairpin loop. Moreover, the two most stable structures (Structure I/-8.1 kcal/mol, and Structure II/-7.9 kcal/mol) incorporate chromophores in two separate hairpin loops. Identification of one of these BrdUs as the cross-linking nucleotide will partially validate these structures and thereby provide important clues as to potential truncation sites in future research.

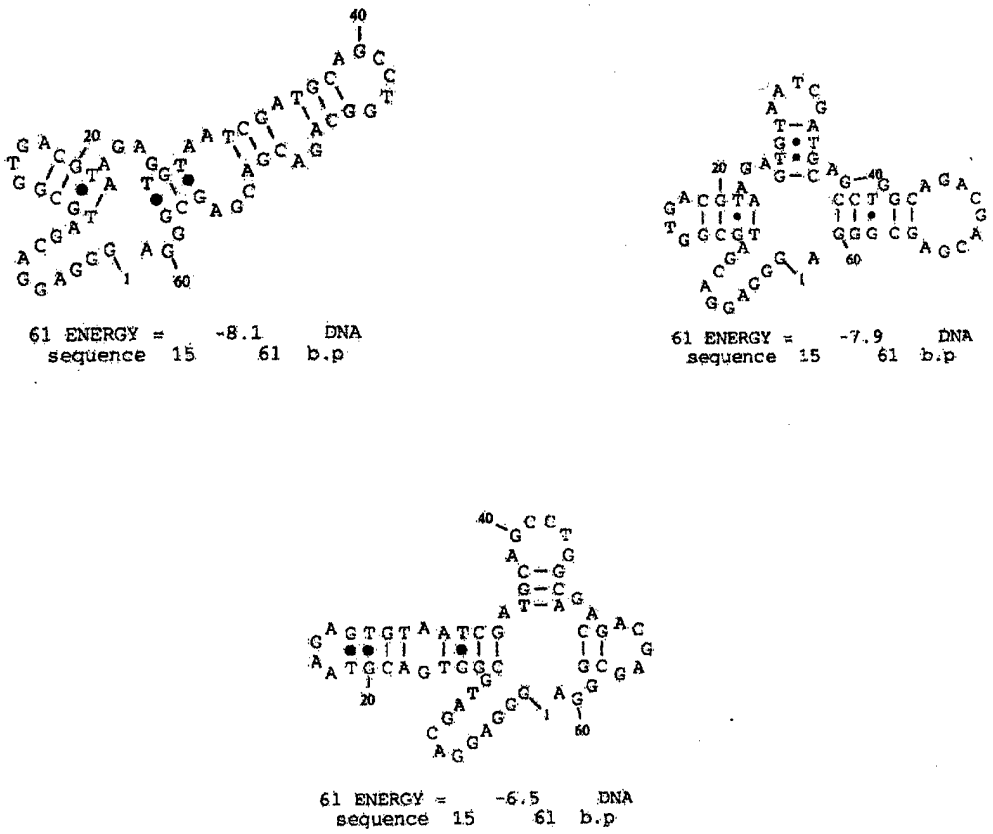


Figure R.2.1. MulFold secondary structures for 06.50.

Although Family II is certainly composed of more than one sequence within each library, the extent of sequence homolgy within the family is high. As mentioned

earlier, multiple copies of numerous Family II aptamers are seen in all three libraries and are therefore subject to the same arguments as 06.50. Since different sequences are available, analysis of covariance is possible and was performed. However, while a highly covariant region was identified in these aptamers which strongly suggested base pairing between nt ~10-16 and nt ~ 42-47, no other information about secondary structure was discernible. As a result, several representative Family II sequences were also folded by the MulFold program and the putative structures are shown below in Figure R.2.2.

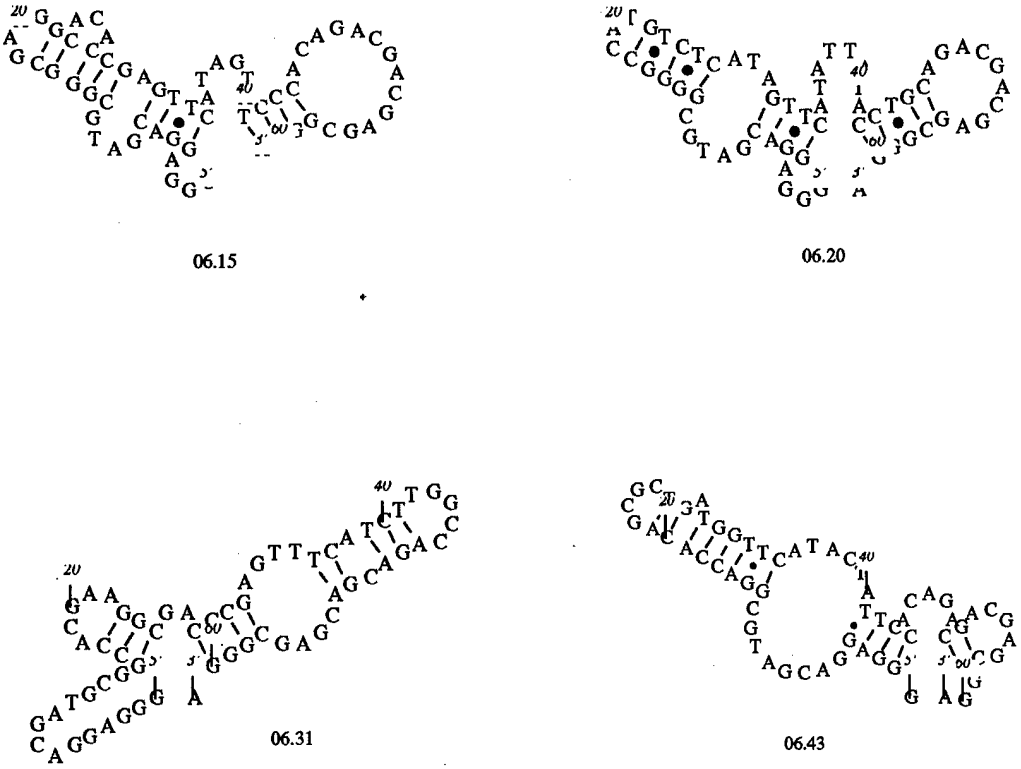


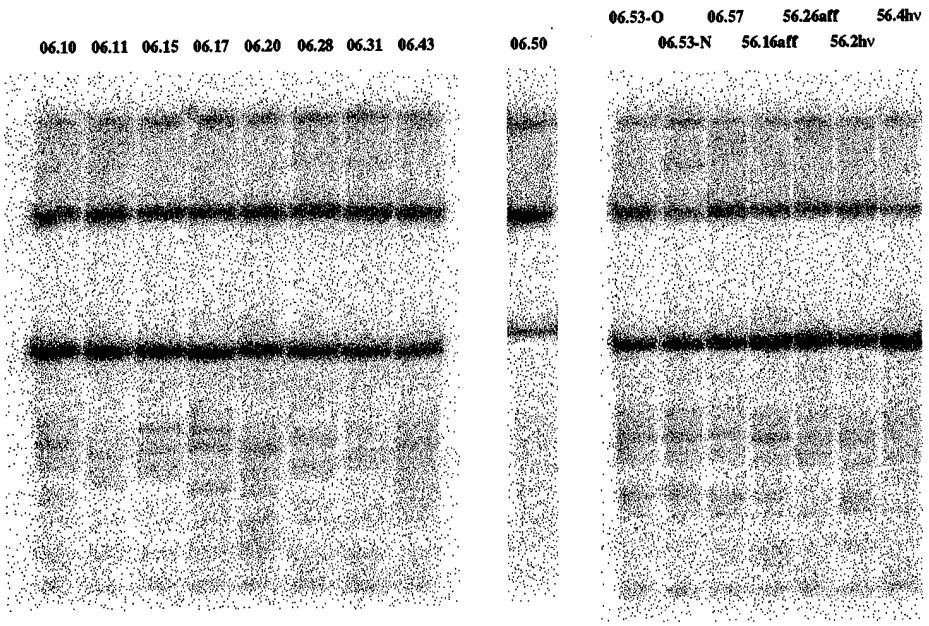
Figure R.2.2. MulFold secondary structures for 0xA/6xP representative aptamers.

One might now conclude that Family II is essentially identical among the three libraries. This would seem to suggest that photochemical pressure is indistinguishable from affinity pressure which opposes fundamental tenets of this research and SELEX dogma. However, closer examination of the 5xA/6xP hν library reveals a subgroup of Family II which does not appear in either of the other two libraries. This subgroup is characterized by exceptionally high BrdU count within the random region. This result might be explained through improved quantum yield afforded by numerous chromophores. Their conspicuous absence from the 0xA/6xP library which employed photochemical pressure might indicate that they do not exhibit particularly high afinity for bFGF. As such, they were not capable of evolving in the face of simultaneous affinity binding pressure.

CHAPTER V: PHOTOCROSS-LINKING AND AFFINITY BINDING SCREENING OF PHOTOSELEX APTAMERS

Once the evolved aptamers had been identified through cloning and sequencing, it was possible to evaluate them to identify those with the highest quantum yield cross-linking and highest affinity binding for bFGF. As illustrated in the previous section, all three libraries exhibited remarkable convergence with several sequences appearing numerous times. To screen these aptamers as systematically as possible, it was decided to first evaluate those sequences appearing multiple times in the 0xA/6xP library for cross-linking yield along with a few sequences which appear uniquely in libraries 5xA/6xP hv or 5xA/6xP aff. It should be noted that the high-content BrdU sequences which evolved in the 5xA/6xP hv library were not chosen for screening as the sequence data for their random regions showed fewer than 30-nt. Those aptamers which showed the highest cross-linking yield were then evaluated for affinity binding.

To screen the chosen aptamers for cross-linking ability, 20,000 cpm total activity of each 5'-^{32}P-radiolabelled oligonucleotide was incubated with 25 nM bFGF (providing a large protein excess over DNA) and irradiated with 1000 pulses of 308 nm light. The reactions were then partitioned by 12% PAGE and evaluated by phosphor-image as seen in Figure R.3.1 on the following page. These results clearly illustrate that 06.50 is a legitimately-evolved aptamer for bFGF. In fact, its superior cross-link yield reaffirms the contention that PhotoSELEX <u>can</u> identify the highest quantum yield cross-linking aptamer from all those capable of cross-linking. These results also confirm again the fundamental edict of any SELEX protocol: evolved aptamers correspond to the applied pressure. Finally, they serve to answer a lingering



Aptamer	*% XL yield*	*Aptamer*	*% XL yield*
06.10	23.2	06.43	28.0
06.11	28.1	06.50	53.6
06.15	24.2	06.53	23.3
06.17	24.4	06.57	15.4
06.20	26.4	5xA/6xP hv	21.9
06.28	26.7	5xA/6xP aff	14.6
06.31	29.9	5xA/6xP aff	21.9

Figure R.3.1. Screening of chosen aptamers for cross-link yield.

question but also pose a new one.

The revelation that 06.50 possesses the highest quantum cross-linking ability of all sequences tested strongly bolsters the claim that PhotoSELEX is a viable combinatorial chemistry technique. The fact that this sequence appears multiple times in each library plainly indicates it enjoys substantial competitive advantages over all other aptamers as discussed in the previous chapter. Moreover, the fact that this aptamer does not share any apparent sequence homology with any other aptamer supports the contention that this may be the one perfect oligonucleotide for cross-linking to bFGF of **all** those which were ever present in any library. This is yet another example of PhotoSELEX's unique ability to effectively evaluate all sequence space and rapidly evolve the best aptamers. This, in conjunction with the decidedly lower cross-link yields for at least one of the 5xA/6xP aff unique aptamers, also serves as another example that the SELEX concept evolves aptamers according to the pressures applied.

In addition, these results help answer a question resulting from the observed cross-link yield of the evolved libraries. Recall from the PhotoSELEX experiments chapter that the maximum cross-link yield measured for the 5xA/6xP hv library was 29.7% and occurred at the cross-link conditions of 50 nM DNA/50 nM bFGF/2000 pulses 308 nm light. However, approximately 7% cross-link yield was measured for the same library at 50 nM DNA/10 nM bFGF/250 pulses 308 nm light. Simply extrapolating this result to the one-to-one condition would seem to indicate that approximately 35% cross-link yield should be observed after 250 pulses. The hypothesized reason for this five-fold excess paradox was that a higher DNA/protein ratio forced more competition between aptamers such that only the highest quantum yielding cross-linkers were allowed to react. When this competition was relaxed,

other lower quantum yield cross-linking aptamers were allowed to react with subsequently lower cross-link yield than anticipated. The measured cross-link yield of 53.6% for 06.50 at protein excess and 1000 pulses of 308 nm light helps substantiate this hypothesis. In this case, the library is composed of one sequence (all of which cross-link with the same quantum yield). Therefore, performing the cross-link reaction at 1:1 or even protein excess gives cross-link yields more consistent with the value expected from linear extrapolation of the cross-link yield observed at DNA excess. The five-fold excess paradox is, therefore, a library phenomenon.

Finally, these results posed an interesting question. All PAGE gels observed thus far suffer from a substantial well-band and a diffuse region of DNA migration above the cross-link band. The lower the DNA/protein ratio, the more substantial the well band and diffuse region. However, gels presented in this section which were used to screen cross-link yield, however, show a substantial reduction (although still present) of these effects. It seemed likely that intermolecular bFGF disulfide bonds (see Introduction on bFGF) leading to one molecule of ssDNA cross-linked to dimers, trimers or even oligomer protein aggregates were the cause of these variable well-band and diffuse, slowly-migrating regions. To test this hypothesis, 20,000 cpm total activity of 06.50 was cross-linked with 25 nM bFGF and 50 nM 06.50 with 20,000 cpm total activity was cross-linked with 50 nM bFGF at 1500 pulses 308 nm light. This time, the reaction was partitioned via a 12% sodium dodecyl sulfate (SDS)-PAGE. SDS-PAGE denatures protein (with SDS) and reduces protein disulfide bonds (with dithiothreitol) to eliminate aggregation. As shown below in Figure R.3.2 on the following page, under these conditions, the cross-link yield of 06.50 as determined by phosphor-image is substantially improved.

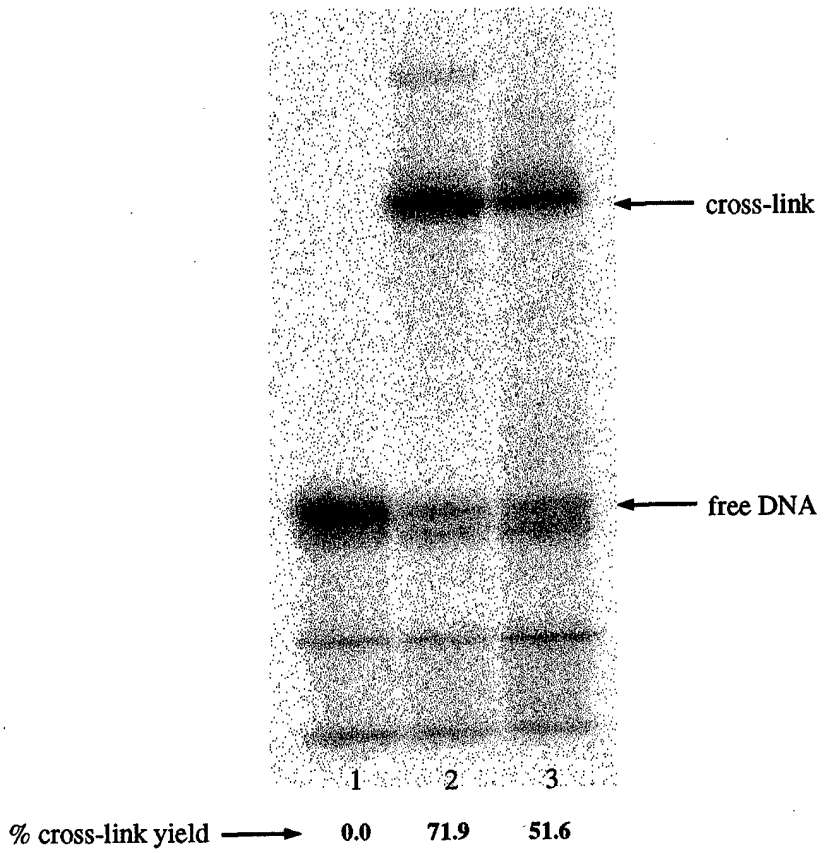
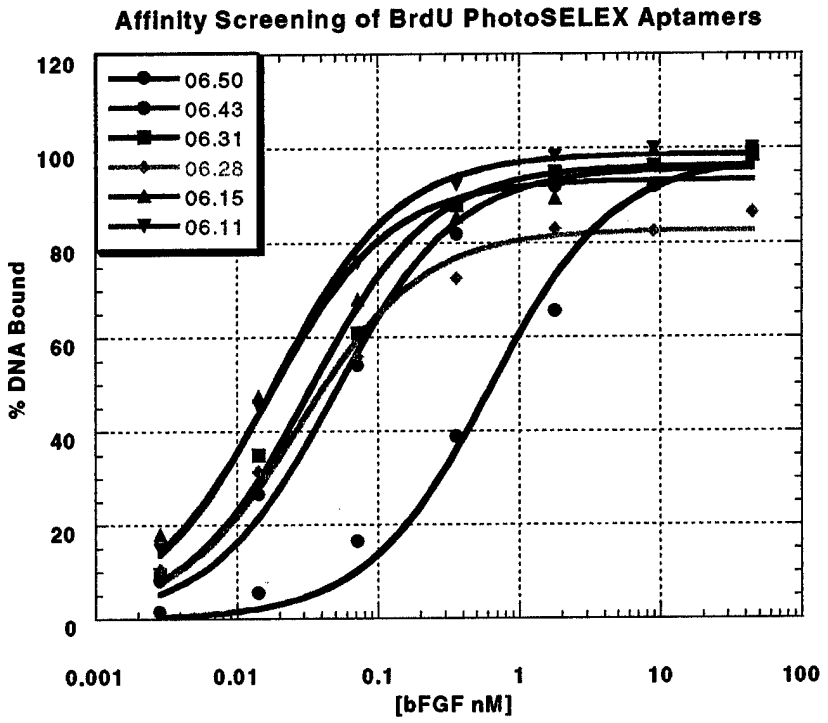


Figure R.3.2. 06.50/bFGF cross-link yield as measured by phosphor-image of a 12% SDS-PAGE partition of the cross-link reaction. (Lane 1= 20,000 cpm 06.50/25 nM bFGF/0 pulses 308 nm light. Lane 2 = 20,000 cpm 06.50/25 nM bFGF/1500 pulses 308 nm light. Lane 3 = 50 nM DNA(20,000 cpm) 06.50/50 nM bFGF/1500 pulses 308 nm light.)

As a result of bFGF aggregation, the previous cross-link yield results for all aptamers have apparently been underreported. For example, counting all the activity which appears in the cross-link band and above on the 12% PAGE for the screening reaction of 06.50 gives a cross-link yield of 74.2%--very nearly the same cross-link yield as measured by 12% SDS-PAGE for the same cross-linking conditions.

Clearly, PhotoSELEX has evolved a truly outstanding photocross-linking aptamer to bFGF.

Based upon the results obtained from assessment of cross-linking ability, six aptamers were selected for further evaluation by measuring their binding affinity for bFGF (06.11, 06.15, 06.28, 06.31, 06.43, and 06.50). These results are shown in Figure R.3.3 below. Clearly, five of these aptamers show exceptionally high affinity



Aptamer	06.11	06.15	06.28	06.31	06.43	06.50	225t3GC
K_d	0.019	0.016	0.028	0.033	0.049	0.56	0.20

Figure R.3.3. Binding affinity screening of six selected PhotoSELEX-evolved aptamers for bFGF (K_ds reported in nM).

for bFGF. In fact, relatively few affinity SELEX-evolved aptamers have been reported with affinities as high as those shown here (Osborne and Ellington, 1997). As a comparison, the binding affinity of a truncated version of the aptamer 225-61 identified by Collins et al. (manuscript in preparation) by post-affinity SELEX modification is also reported above. It is interesting that of these aptamers the best cross-linker has the lowest affinity (06.50) while the poorest cross-linker has the highest affinity (06.15). While it is not within the purview of this research to ascertain any particular relationship between cross-linking and affinity binding (other than to identify those sequences which possess both characteristics to the greatest degree), these results do hint that highest quantum yield cross-linking may require specific chromophore/amino acid orientations which are not necessarily consistent with conditions required for highest binding affinity. Nevertheless, a fundamental premise of this research has been borne out: high quantum yield cross-linking and high binding affinity are not disjoint characteristics and can be selected for simultaneously from a library of modified oligonucleotides. Based upon the results presented in this section, two aptamers, 06.15 and 06.50 and their associated bFGF cross-linked adducts were more fully characterized.

CHAPTER VI: EVALUATION OF THE DIAGNOSTIC POTENTIAL OF 06.15 AND 06.50

A fundamental motivation for evaluating the ability of PhotoSELEX to identify high-affinity binding/high quantum yield cross-linking aptamers for protein targets is for its application as an identifier of diagnostic molecules. Effective diagnostic tools, regardless as to their mode of operation, share two essential qualities. First, they are sensitive to the target or condition for which they test. The more sensitive a diagnostic test, the earlier a condition can be detected and the easier it can be corrected. Second, they are selective for the target or condition for which they test. Similar targets or conditions should be readily distinguished from signal. In the vernacular of medical diagnostics, these two characteristics are referred to as minimizing the probability of false negatives and false positives, respectively. As such, it was necessary to assess the sensitivity and specificity of 06.15 and 06.50 for bFGF.

One aspect of the sensitivity of 06.15 and 06.50 to bFGF, their binding affinity, was reported in the previous section. It is beneficial, however, to put these affinities ($\approx$ 16 pM for 06.15 and $\approx$ 560 pM for 06.50) into perspective. As will be discussed in detail later, the current state of the art for diagnostic testing for proteins and antigens is the immunoassay. The sensitivity of the immunoassay depends primarily on the binding affinity of a monoclonal antibody for the antigen. In the case of bFGF, the binding affinity of the monoclonal antibody permits identification of 1.0 ng/well of bFGF when employing an antibody concentration of 0.5-1.0 μg/ml (R&D catalog, 1997). If the K_ds of the PhotoSELEX-evolved aptamers allow detection of bFGF at equal or lesser concentrations than these, then PhotoSELEX would seem capable of

competing with current antibody diagnostic technology.

Another aspect of the sensitivity of the evolved aptamers for bFGF which has not been evaluated is essentially the photocross-linking analog to the thermal binding curve. The objective of this experiment is to measure how cross-link yield correlates with protein concentration. As such, this offers a direct assessment of the sensitivity of 06.15 and 06.50 for bFGF as they might be employed in a diagnostic system. In this experiment, various bFGF concentrations were prepared by serial dilution with 1x PBS (just as when performing a binding curve). To each concentration of bFGF (including a sample with 0 nM [bFGF]) a quantified activity of 5'-^{32}P-radiolabelled aptamer was added (again, just as when performing a binding curve). However, rather than measuring signal as sequestered by nitrocellulose filter binding, the aptamer/protein mixtures are irradiated at the photocross-linking condition (308 nm light/1500 pulses) and the result of each reaction is analyzed by SDS-PAGE. The lower the bFGF concentration that yields a cross-link signal, the higher the sensitivity of the aptamer for bFGF. The results of this experiment are shown in Figure R.4.1 on the following page. In both cases, a cross-link signal can be observed at bFGF concentrations as low as 0.0032 nM using a total activity of 20,000 cpm in each lane and 0.5 h phosphor-image exposure. This is exceptional sensitivity and corresponds to a detection capability of 0.00115 ng or 0.058 ppt of bFGF—nearly three orders of magnitude greater sensitivity than the antibody for bFGF using ELISA detection. Confidence in this result is enhanced since a distinct cross-link signal difference is seen between 0.0032 nM and 0.0 nM [bFGF]. Presumably, even lower concentrations could be detected via SDS-PAGE by increasing the total activity and exposure time.

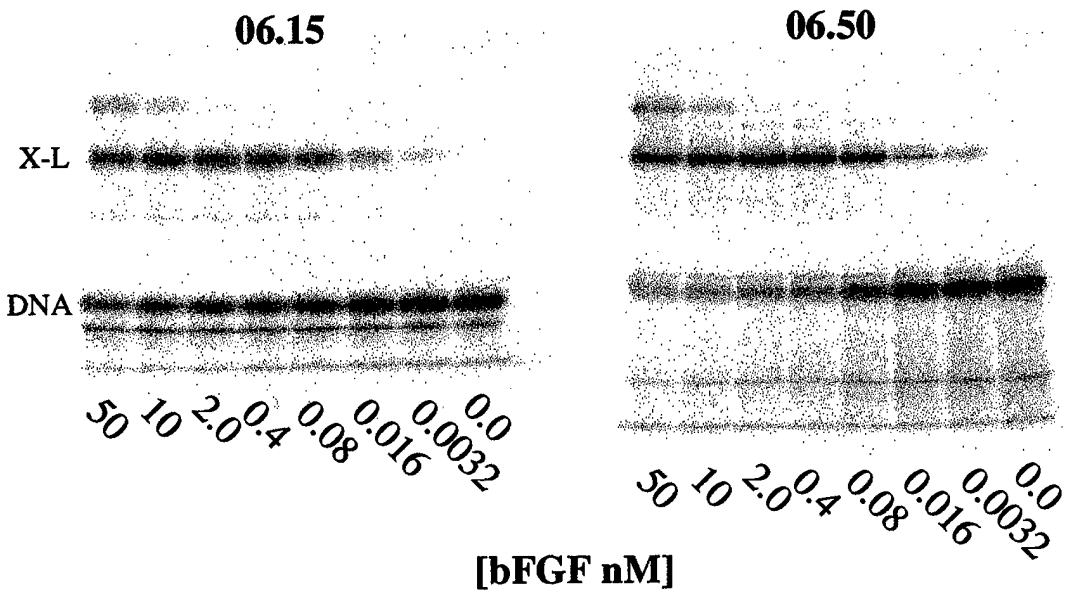


Figure R.4.1. Sensitivity of PhotoSELEX-evolved aptamers for bFGF as detected by SDS-PAGE analysis of cross-linked adduct formed in PBS.

It is interesting and somewhat counterintuitive that while 06.15 exhibits substantially better affinity binding for bFGF than does 06.50, they both show cross-link signal at 0.0032 nM [bFGF]. In fact, the cross-link signal at 0.0032 nM [bFGF] is slightly greater for 06.50 than for 06.15. Evidently, 06.50's superior quantum yield cross-linking ability more than compensates for its inferior binding affinity. This can be rationalized through examination of the dissociation constant, K_d where $K_d = k_{off}/k_{on}$ (k_{off} = rate of DNA/protein complex dissociation, k_{on} = rate of DNA/protein complex association). As such, the lifetime of the 06.15/bFGF affinity adduct is

substantially longer than that of the 06.50/bFGF affinity adduct. However, the time for the photochemical reaction is much shorter than the lifetime of either affinity adduct and will trap the affinity adduct a given percentage of the time based upon the quantum yield. Therefore, while both parameters are important in the identification of photocovalent diganostic molecules, this result seems to indicate that quantum yield cross-linking is the critical capability. Nevertheless, both 06.15 and 06.50 exhibit tremendous sensititivity to bFGF as measured by SDS-PAGE analysis of the cross-link signal. It was now necessary to assess the specifity of these aptamers for bFGF.

The specificity of 06.15 and 06.50 for bFGF was determined in two experiments. In the first experiment, the binding affinity of these aptamers was measured for two other members of the heparin-protein binding family: vascular endothelial growth factor (VEGF) and platelet derived growth factor (PDGF). As before, this was measured by nitrocellulose filter binding with the results indicated in Table.R.4.1. These results show that both 06.15 and 06.50 readily distinguish between even consanguineous proteins based on affinity binding.

Table R.4.1. Specificity of 06.15 and 06.50 for bFGF vis a vis VEGF and PDGF based on normalized binding affinity.

SPECIFICITY OF 06.15 AND 06.50 AFFINITY ADDUCTS			
Affinity Adduct	**bFGF**	**VEGF**	**PDGF**
06.15	1	<0.001	<0.001
06.50	1	<0.001	<0.001

In the second experiment to determine specificity, the cross-link signal for 06.15 and 06.50 was again measured via SDS-PAGE. Unlike the reaction conditions for determination of sensitivity, however, the specificity cross-link test used 1x PBS with 10% serum as the cross-linking medium. Since serum contains an immense array of proteins, it is likely that nucleic acid and heparin binding proteins are present. Therefore, any cross-link signal in the absence of bFGF would immediately call into question the cross-linking specificity of these aptamers. As can be seen in Figure R.4.2 below, however, neither 06.15 nor 06.50 yield any cross-link signal in the absence of bFGF.

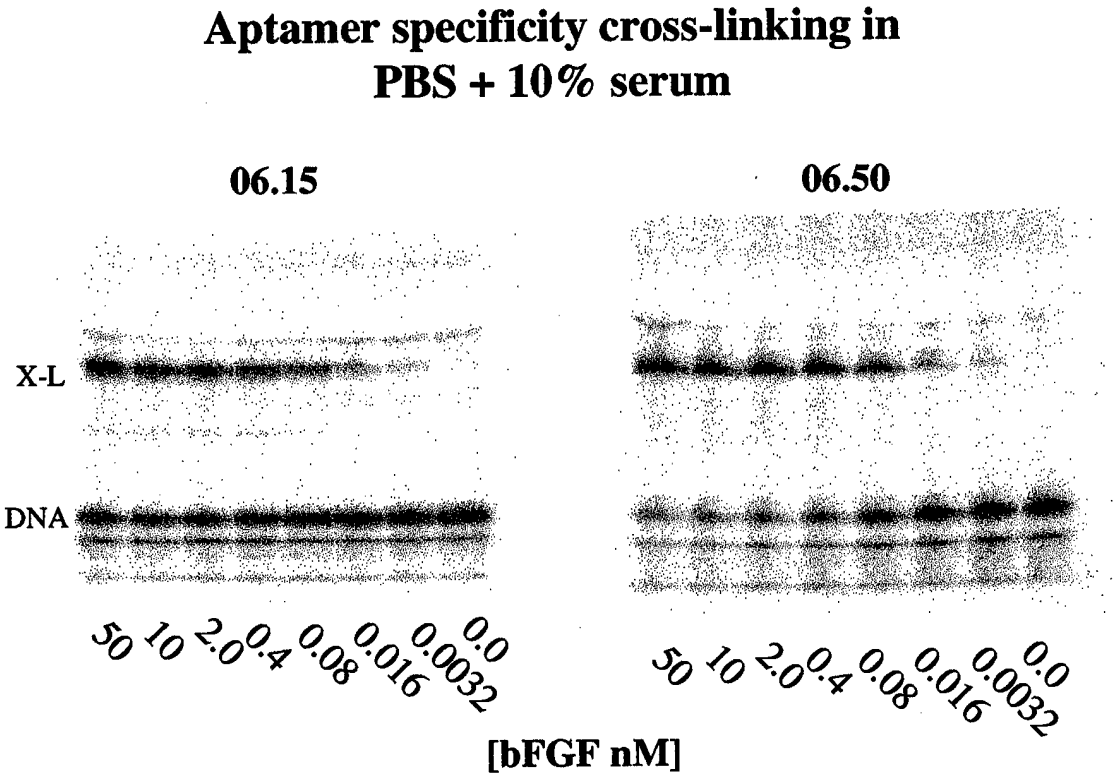


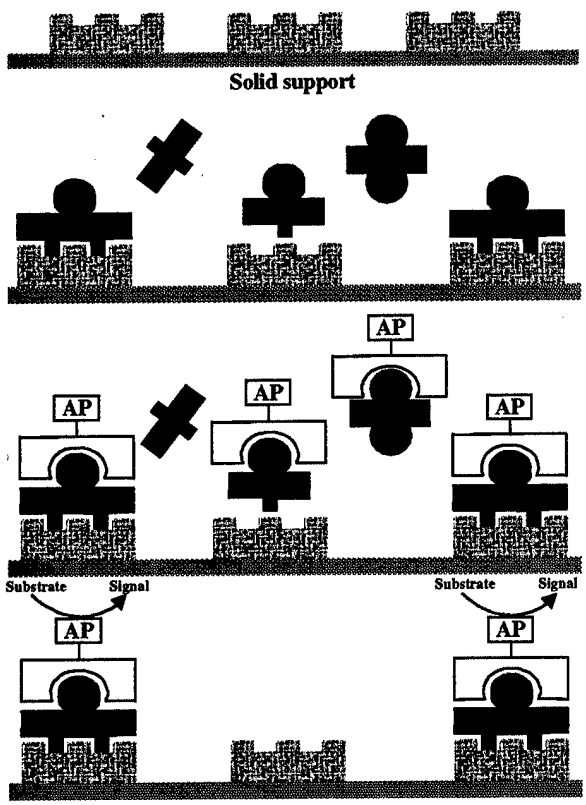
Figure R.4.2. Specificity of PhotoSELEX-evolved aptamers for bFGF as analyzed by SDS-PAGE partition of cross-linked adducts formed in PBS/10% serum.

This experiment also serves to illustrate the feasibility of these aptamers for use in a diagnostic system. Since identification of protein markers or specific antigens from a patient will be made from a serum extract, the observation that the sensitivity of these aptamers is not reduced in a 10% serum cross-link environment is very encouraging. It should be noted, however, that the slightly greater sensitivity to bFGF exhibited by 06.50 in the PBS environment is lost in the 10% serum environment. Careful examination shows that the cross-link signal for 06.50 is nearly halved in the 10% serum environment for 0.0032 nM [bFGF] compared to the signal detected for 0.0032 nM [bFGF] in the PBS environment. Conversely, virtually no change in cross-link yield is observed for 06.15 under the same conditions. Here then, is strong evidence in support of the importance of target affinity. Evidently, as nonspecific interference to cross-linking increases, so does the necessity to aggressively bind to the target in order to facilitate photocovalent bond formation. These results clearly call for further investigation in order to identify the very best photocross-linking aptamer for use in a diagnostic system.

Despite the slight ambiguity as to the relative importance between aptamer cross-linking quantum yield and binding affinity, it appeared both 06.15 and 06.50 had tremendous potential to be successfully employed in a diagnostic system for bFGF. As a result, it seemed useful to dedicate some discussion to the practical design of a practical diagnostic replacement to current immunoassay systems.

Immunochemical techniques for the detection of proteins, antigens and analytes were first described in the early 1970s (Engvall, et al., 1972, Hoffman, 1973, Ljungstrom, et al., 1974, Engvall, et al., 1971). The major technique to evolve from

these orignal investigations is the enzyme-linked immunosorbant assay (ELISA). While many different ELISA techniques now exist (e.g. sandwich immunoassay, monoclonal screening, and competitive assays among others) fundamental characteristics apply to all. Since the current standard for detection of large molecules like proteins is the sandwich immunoassay, this will be examined exclusively and is illustrated in Figure R.4.3 below.



Step 1: Immobilization of the capture antibody on a solid support.

Step 2: Incubation of sample containing analyte molecules of interest and other molecules.

Step 3: Addition of the detector antibody covalently bound to alkaline phosphatase (AP) which recognizes a different epitope of the analyte molecule.

Step 4: Irrigation of the assembly and addition of AP substrate yield immunochemical signal for desired analyte.

Figure R.4.3. The ELISA sandwich immunosorbant assay (after Voet and Voet, 1995).

In the direct sandwich assay, an antibody raised against the protein or analyte

by the immune system of an animal is immobilized on a solid support. This so-called capture antibody must recognize a specific epitope of the analyte molecule. However, the immune system produces many different antibodies (polyclonal antibodies) which typically bind different epitopes of a macromolecule. To achieve a monoclonal antibody, a cell producing the desired antibody is fused with a myeloma cell to yield a hybridoma which can produce virtally unlimited quantities of the antibody. Once the capture antibody is immobilized, the system is incubated with the sample (Voet and Voet, 1995).

The capture antibody binds the analyte molecule from the sample after the sample is introduced and incubated on the solid support. As a result of binding a specific epitope of the analyte molecule, the analyte assumes a unique and homogeneous orientation with respect to the solid surface. Next, a detector antibody which is covalently attached to alkaline phosphatase is added. The detector antibody is obtained in precisely the same manner as the capture antibody, but it must recognize a different epitope (ideally one made available by the analyte's current orientation). Finally, the sandwiched array is washed and a substrate for the enzyme such as a 1,2-dioxetane is added. When alkaline phosphatase acts on this substrate, a chemiluminescent signal is generated. An advantage of the direct sandwich immunoassay is signal amplification which results from the presence of two antibodies in the sandwich array.

While the sandwich ELISA has served as a reliable mainstay for protein and antigen detection, it is a costly, time consuming and labor intensive technique. There are currently efforts underway to expand the classic 96-well microtitre plate to much

larger capacity as well as attempts to automate the process. Despite these investigations, there is currently no system available which permits the detection of a wide array of important marker proteins in a patient from a single, small sample of patient serum. PhotoSELEX may permit the development of just such a system.

In the PhotoSELEX diagnostic system, illustrated in Figure R.4.4 below,

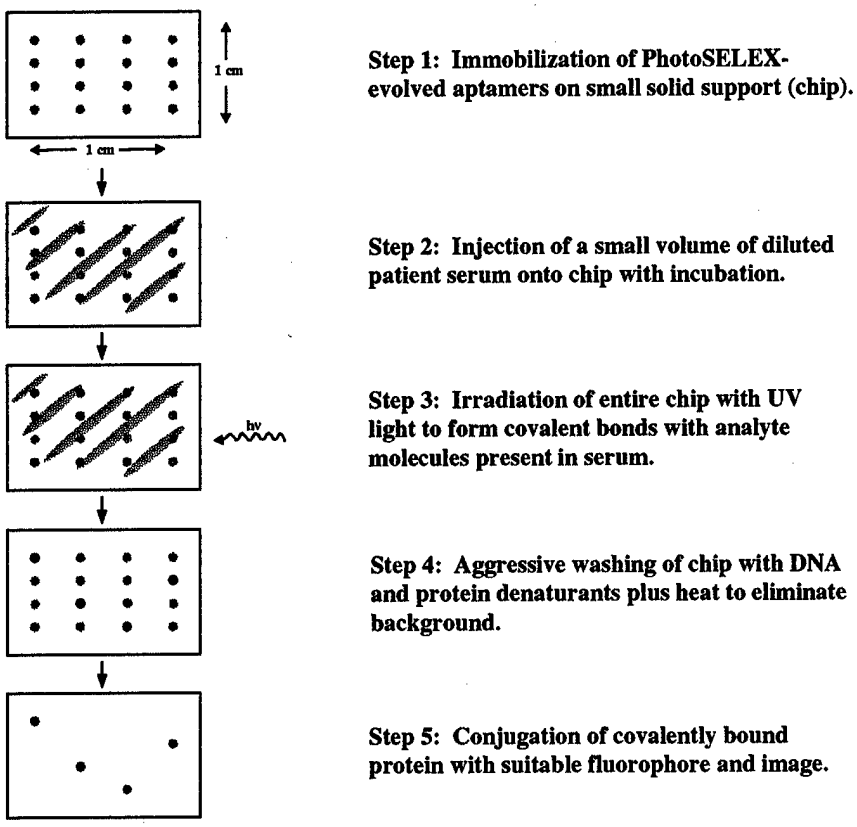


Figure R.4.4. The PhotoSELEX fluorescent assay concept.

PhotoSELEX-evolved apatamers would be generated for many different important marker proteins. Once evolved, these aptamers would be immobilized on a solid surface somewhat like the capture antibody of the ELISA sandwich technique.

However, rather than immobilizing just one type of aptamer somewhat randomly within a well of a microtitre plate, the various evolved aptamers would be fixed on a very small solid support in a specific pattern such as in a grid. The ability to deposit DNA precisely on a solid support has been demonstrated and used to identify active genes in fibroblasts in response to serum introduction (Vishwanath, et al, 1999). These researchers were able to assemble a microarray of cDNA representing about 8600 human genes. They subsequently prepared flourescent-labelled cDNA from fibroblast mRNA and thereby determined gene expression. In addition, the company Affymetrix currently holds certain patent rights to its technique for binding short oligonucleotides representing human genes to a solid support. Presumably, similar techniques could be used to bind PhotoSELEX-evolved aptamers even though they are relatively large oligonucleotides.

With the aptamers bound to the solid support (chip), a suitably diluted serum sample could then be applied to the chip with each aptamer forming an affinity association with its target molecule. After appropriate incubation, the solid support would be irradiated with UV light permitting covalent bond formation between aptamer and target. At this point, the support could be extensively washed using urea, formamide, SDS, DTT and heat. Due to the covalent bond formed between aptamer and protein, any spurious and nonspecific associations can be aggressively eliminated with a significant reduction in background. Finally, the entire solid support would be reacted with a fluorescent dye specific to certain amino acids such as Molecular Probes' Oregon Green succinimidyl ester. This dye yields a larger number of fluorophore conjugates/protein with subsequently greater signal than fluorescein

isothiocyanate. The entire chip would then be examined by a suitable instrument such as a microscrope with a charge-coupled device.

In the PhotoSELEX-based diagnostic system, hundreds of analyte molecules could be tested for with a single patient serum sample. This compares very favorably to the ELISA sandwhich immunoassay system which is currently used to test for a single analyte molecule from numerous patients at a time. Based on this cursory analysis, it seems the PhotoSELEX-based diagnostic system concept can achieve extremely high throughput, would be easily automated, and would require minimal sample--all while achieving extraordinary sensitivity and specificity.

CHAPTER VII: EDMAN DEGRADATION AND MASS SPECTROMETRIC DETERMINATION OF THE BFGF CROSS-LINKING AMINO ACID

An important aspect in the characterization of any oligonucleotide/protein cross-link system is the identification of the protein's cross-linking amino acid. As mentioned in the Introduction, in many cases the primary motivation for forming the photocovalent cross-link bond is to identify the point of contact between the oligonucleotide and the protein. While several techniques have been considered to locate the cross-linking amino acid, the primary method used to date is Edman degredation of the protein after formation of the covalent bond.

The fundamental premise of Edman degredation identification of a protein's cross-linking amino acid is that this amino acid will be absent from the protein's known sequence. For example, one might envision sequencing a given protein starting at the amino terminus with phenylisothiocyanate and trifluoroacetic acid through to the carboxy terminus. After treatment with aqueous acid, each phenylthiodantoin-amino acid would be detected chromatographically by HPLC. Finally, this sequence would be compared to the protein's known sequence. The amino acid which fails to be detected will then correspond to the cross-linking amino acid (the cross-linking amino acid will be insoluble in organic medium due to its cross-link to the oligonucleotide). In practice, however, Edman degradation protein sequencing is limited to 40 to 60-N-terminal residues as a result of side reactions, incomplete reactions and peptide loss. Also, since as little as one picomole of any individual amino acid may be chromatographically detected, Edman degredation may reliably sequence approximately 10 pmol of a 40-mer peptide.

In order to produce sufficient cross-link material of tractable peptide length

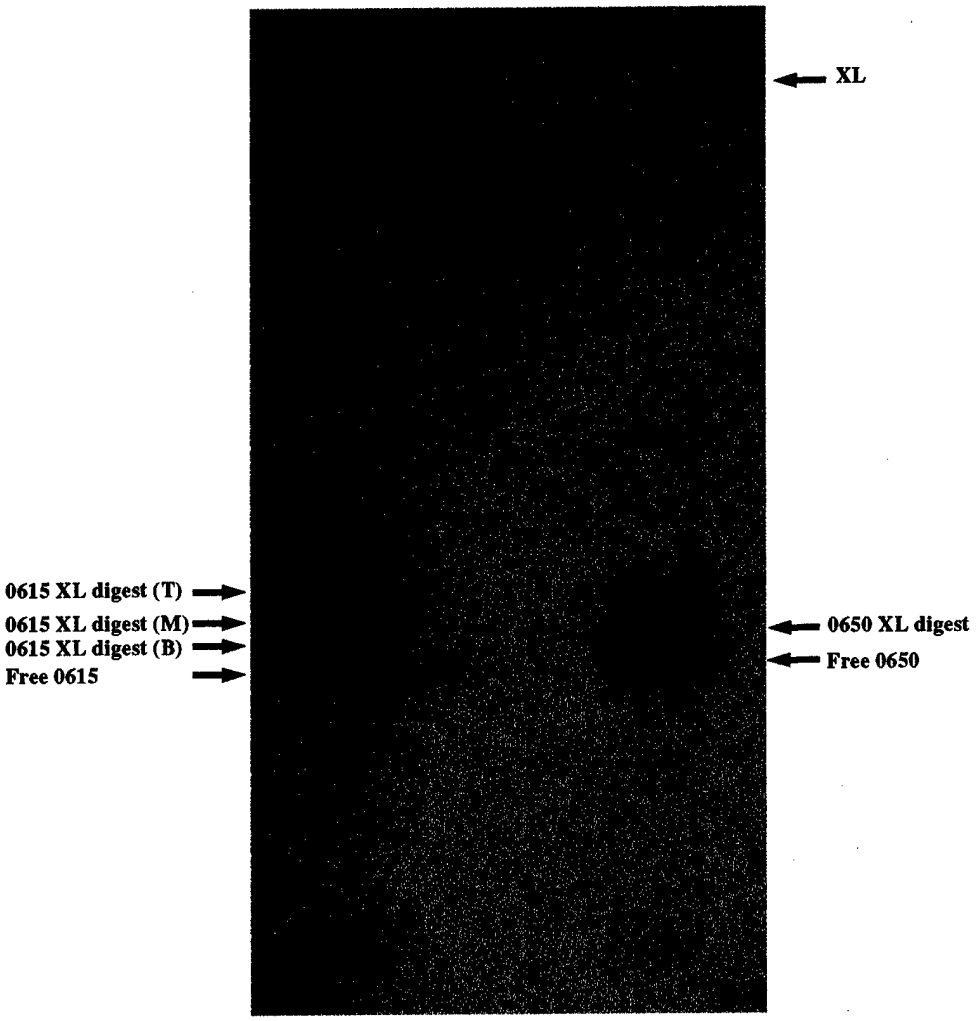


Figure R.5.1. Denaturing PAGE of trypsin-digested 06.15 and 06.50/bFGF cross-linked adducts.

for Edman degradation, 850 pmol with 1,000 cpm/μL specific activity 5'-^{32}P-labelled 06.50 and 06.15 were cross-linked with 283.3 pmol of bFGF in a 1mL reaction volume using 1500 pulses of 308 nm light. This was subsequently digested with a large excess of trypsin, extracted with phenol/chloroform, and partitioned by denaturing PAGE; the result of this procedure is shown in Figure R.5.1 above.

As can be clearly seen, no band appears in the standard cross-link migration region. However, for both aptamers, bands which migrate slightly slower than free DNA do appear. These bands correspond to bFGF tryptic fragments which include the cross-linking amino acid covalently bound to full length 06.50 and 06.15. It is interesting that multiple slightly-slower migrating bands appear for both sequences. In the case of 06.15, this effect is particularly apparent. These results can be interpreted as a consequence of the bulky oligonucleotide interfering with trypsin's ability to cleave efficiently after all Lys/Arg residues. This observation is consistent with the conclusion of Gospodarovicz and Cheng (1986) and Saksela and coworkers (1988) who observed that heparin protects bFGF against enzymatic digestion.

As a result of this somewhat varied digestion, the large 06.50/bFGF tryptic fragment band, and all three of the 06.15/bFGF tryptic fragment bands were excised and eluted from the gel with an estimate of the quantity of recovered material based upon specific activity. Finally, approximately 60 pmol of 06.50/bFGF tryptic fragment and 30 pmol of each 06.15/bFGF tryptic fragment were submitted for Edman degradation sequencing with the results shown below.

06.50

Arg-Thr-Gly-Gln-___-Lys-Leu-Gly-Ser-Lys

06.15

Gln-___-Lys-Leu-???-Ser (Top)

Gln-___-Lys-Leu (Middle)

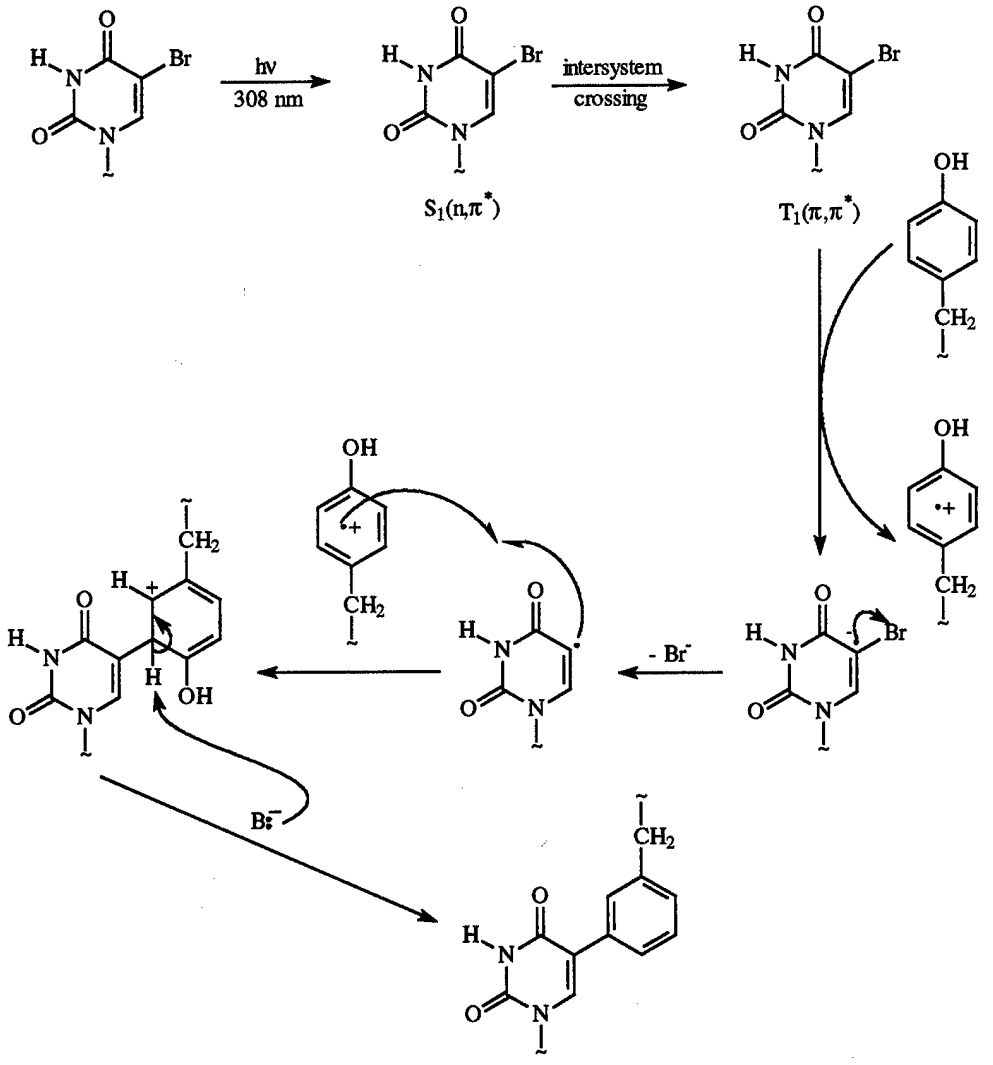
Gln-___-Lys (Bottom)

Figure R.5.2. Edman degradation sequencing of aptamer cross-linked/tryptic bFGF.

It is difficult to over emphasize the significance of these results. First, they again substantiate a fundamental premise of SELEX in general and PhotoSELEX in particular: the photochemical pressure has identified the most fascile photocross-linking amino acid. Second, they show that PhotoSELEX has identified aptamers which bind the most crucial component of bFGF's heparin-binding epitope.

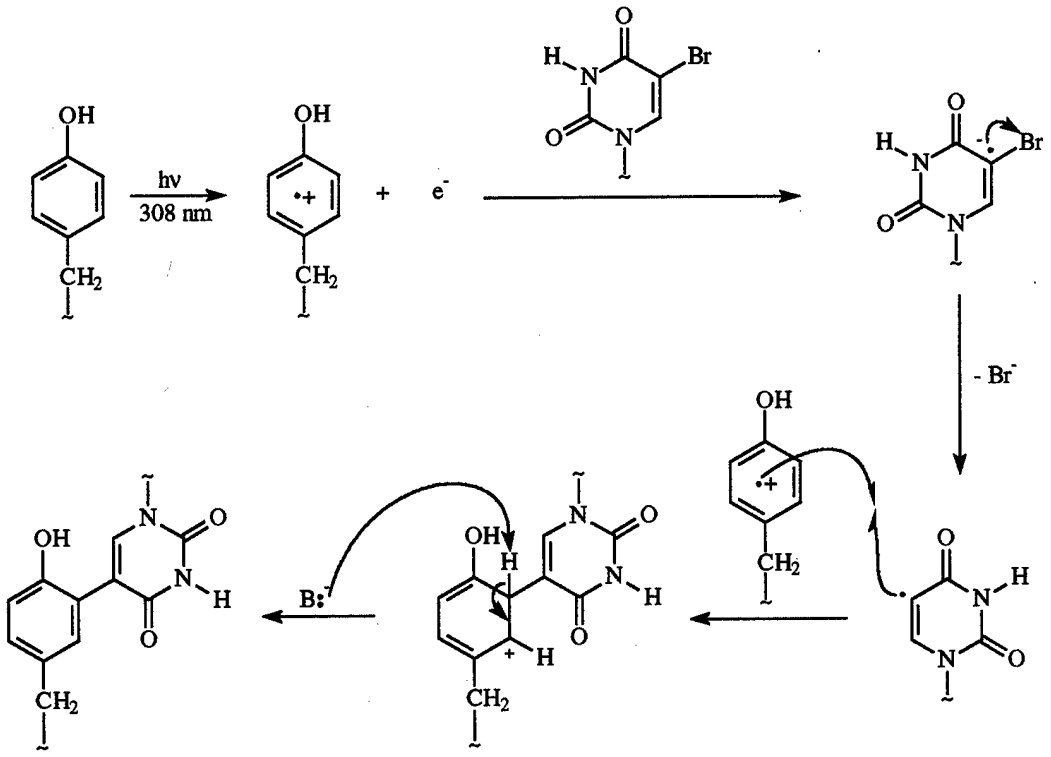
To begin, PhotoSELEX has evolved aptamers which cross-link with tyrosine-the most photoactive amino acid available. As discussed in the review of 5-bromo-

Scheme R.5.1. Excitation of BrdU preceding photocross-linking to tyrosine.



2'-deoxyuridine's photochemistry in the Introduction, the aromatic amino acids including histidine, tryptophan, phenylalanine and tyrosine posses the greatest ability of all the amino acids to cross-link with BrdU. This is a result of the relatively low ionization potential of these residues. Although it is still uncertain whether it is excitation of BrdU (first to S_1 with subsequent intersystem crossing to T_1) or excitation of tyrosine with subsequent ejection of an electron which occurs first, both mechanisms ultimately involve the transfer of a single electron from tryrosine to BrdU. This forms the radical ion pair which promotes bromide to leave to create the uracilyl radical. This is followed by radical combination to form the covalent bond

Scheme R.5.2. Excitation of tyrosine preceding photocross-linking to BrdU.



and finally proton abstraction to give the neutral adduct. These photochemical routes are shown in Schemes R.5.1 and R.5.2, respectively, on the preceding two pages.

Secondly, PhotoSELEX has evolved aptamers which bind to the most critical region of bFGF's heparin-binding epitope. Although 06.15 and 06.50 do not seem to share any apparent sequence homology, they both cross-link to Tyr133(133) of $bFGF_{155}$ whose position is shown in Figure R.5.3 on the following page. As described in the Introduction, Moy and coworkers (1996), have identified a helix-like conformation within the known heparin-binding epitope of bFGF which they surmized may play a crucial role in heparin recognition. It is particularly interesting to note that as this helix includes residues 131-136, PhotoSELEX has evolved aptamers which covalently bind at the heart of the heparin binding site. Eriksson and coworkers (1993) also contend that substantial evidence implicates Lys125(134) as the key to heparin binding. While these results are very positive in themselves, the ability of PhotoSELEX to identify such high quantum yield cross-linking aptamers for a protein target presented a new opportunity to explore the ability of mass spectrometry to sequence and identify peptides with post-translational modifications.

Over the past 30 years mass spectrometry has emerged as an exceptionally useful tool in macromolecular characterization. The ability of mass spectrometry to sequence simple peptides was first demonstrated in the early 1970s (McLafferty et al., 1973, McLafferty et al., 1973b). Somewhat more recent advancements in mass spectrometric methods have permitted sequencing of polypeptides providing an attractive alternative to Edman degredation which cannot sequence peptides with post-translational modifications (e.g. blocked N-termini and glycosylated, phosphorylated or acetylated residues). In particular, fast atom bombardment (FAB)

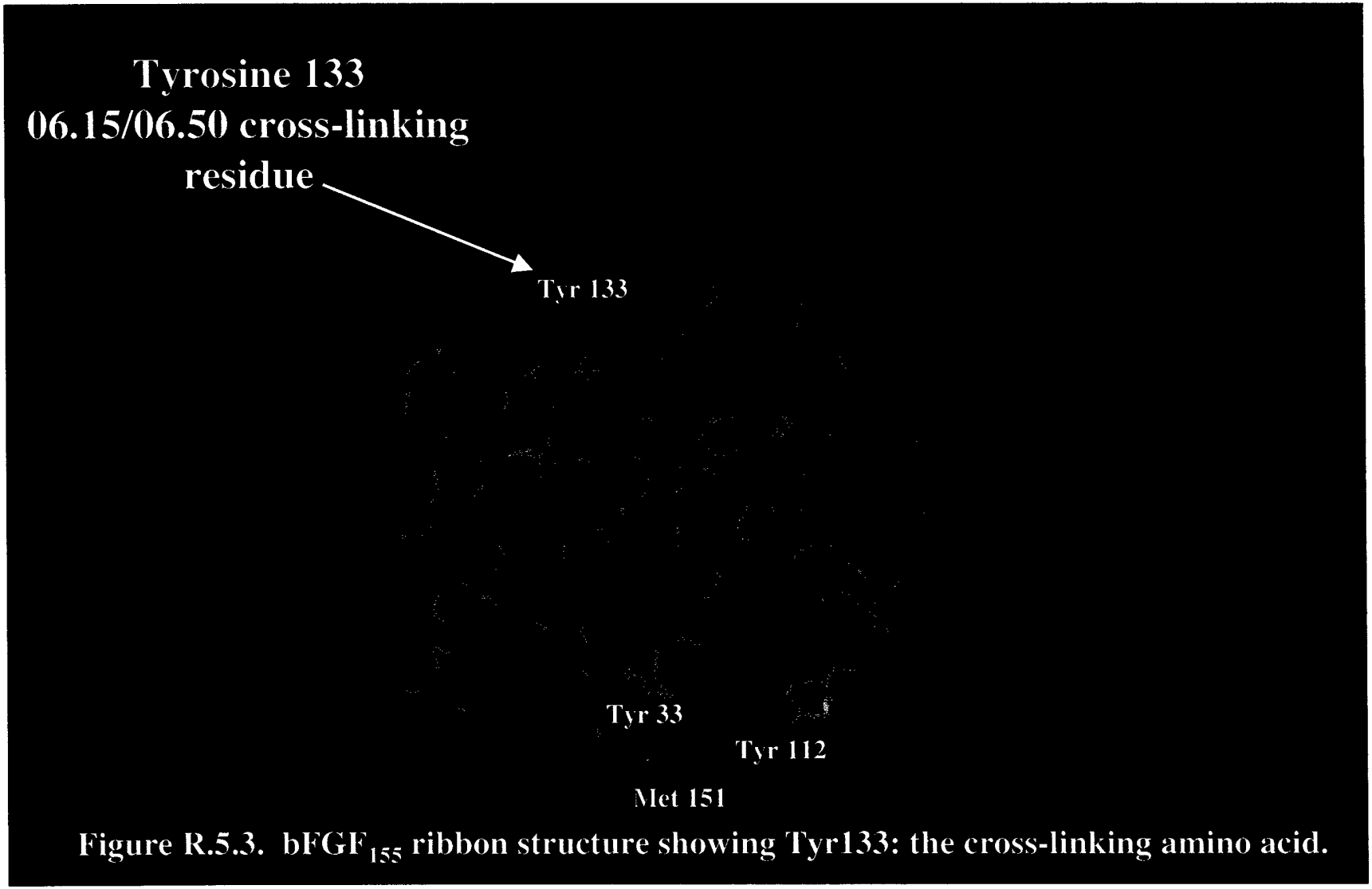


Figure R.5.3. $bFGF_{155}$ ribbon structure showing Tyr133: the cross-linking amino acid.

with tandem mass spectrometry (MS-MS) has evolved as a successful mass spectrometric method for polypeptide sequencing. In this technique, the polypeptide is first dissolved in a solvent of low volatility such as glycerol and subsequently bombarded with a beam of argon or xenon atoms. Bombardment of the glycerol/peptide solution yields $(M+H)^+$ ions which can be detected and selected with the first mass spectrometer from other peptides or unwanted contaminants which may be present. The selected ion can then be fragmented by He atoms (or other inert atoms) and analyzed by the second mass spectrometer. The key to the success of FAB/MS-MS (as well as more advanced mass spectrometric methods used in this research) is that the polypeptide will fragment in known ways from both the N and C-termini to yield fragmentation patterns which can be deciphered as an unique amino acid sequence. Since correct spectral analysis is so important to sequence identification, a very brief discussion of MS polypeptide fragmentation nomenclature is warranted.

Roepstoff and Fohlman (1984) with subsequent modification by Biemann (1988) established the nomenclature currently used for peptide fragmentation ions. The assignment of a peptide fragment ion is based upon whether the fragment results from amide-backbone bond or residue side chain bond cleavage and whether charge retention occurs with either the N-terminal or C-terminal amino acid fragment. Figure R.5.4 on the following page illustrates the amino acid fragment ions which can be generated in MS from backbone bond cleavage and the associated nomenclature. Although FAB/MS-MS readily generates fragment ions resulting from side chain bond cleavage, more advanced techniques employed in this research generate backbone bond cleavage almost exclusively.

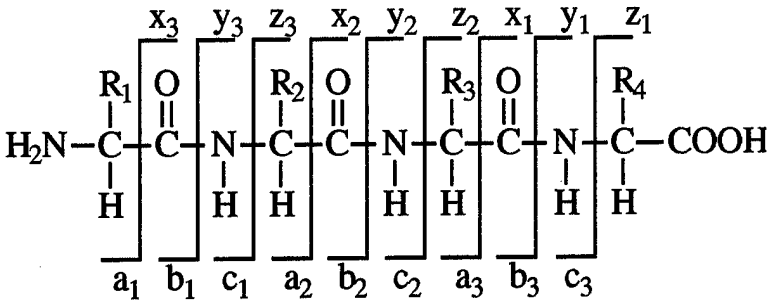


Figure R.5.4. Definition of the amino acid fragment ions resulting from backbone bond cleavage.

FAB/MS-MS enjoys two particular advantages over Edman degradation. First, as mentioned above, mass spectrometry can sequence peptides with post-translational modifications. Recall that identification of Tyr133 as the cross-linking amino acid by Edman degredation was made by omission and thus depends on a prior knowledge of the peptide sequence. Second, mass spectrometric peptide sequencing can be completely automated and computerized making the process much faster than Edman degradation. The significant disadvantage associated with FAB/MS-MS is reduced sensitivity. Since it is difficult and expensive to produce nanomolar quantities of trypsin digested cross-linked nucleoprotein adduct, reduced sensitivity is a substantial drawback to any sequencing methodology. Fortunately, more recent advances in mass spectrometric techniques have significantly enhanced sensitivity.

Electrospray mass spectrometry has recently emerged as a very effective method for sequencing peptides with post-translational modifications. In this technique, an organic/aqueous solvent mixture (including the analyte) is injected into a high potential field through a very small gauge needle. This atomization generates charged droplets which are accelerated toward the instrument's analyzer. This

acceleration phase takes place in a vacuum which promotes evaporation of the droplet's solvent. Evaporation yields smaller, less stable charged droplets from which molecular ions are subsequently emitted. Finally, these ions are focused and accelerated by quadropole analyzers which separate the ions according to m/z. Electrospray mass spectrometry (ESMS) is unique in that it doesn't require fragmentation to generate ions. As a result, substantially more molecular ions are generated for a given sample with concomitant enhancement in sensitivity. Finally, the m/z ions of interest can be selected and fragmented by the MS-MS method as described above for FAB/MS-MS. Thus, ESMS utilizes the inherent advantages of FAB/MS-MS while eliminating its lack of sensitivity.

Advanced mass spectral techniques have been successfully employed to characterize certain aspects of covalently bound oligonucleotide/protein adducts. For instance, several research groups have shown that matrix assisted laser desporption ionization (MALDI) and HPLC/electrospray ionization (ESI) MS can be used to establish the molecular mass of proteins cross-linked to oligonucleotides (Jensen et al., 1993; Bennett et al., 1994; Jensen et. al., 1994; Connor et al., 1998b; Wong et al., 1998). In addition, sequencing by collision-induced dissociation (CID) of model peptides cross-linked to a homogeneous oligonucleotide T_6 and of peptides cross-linked to mono or dinucleotides has been demonstrated (Jensen et al., 1996; Connor et al., 1998a). A cross-link has also been characterized with MALDI mass spectrometry coupled with N-terminal microsequencing and alkaline RNA hydrolysis (Urlaub et al., 1997). Although these cases show the potential for a mass spectral characterization of protein-nucleic acid cross-links, MS sequencing of a cross-linked high affinity protein-nucleic acid complex by CID has not been previously reported.

To evaluate the potential for ESMS sequencing to identify the cross-linking amino acid in bFGF, trypsin digested 06.50/bFGF cross-linked adduct was prepared precisely as described for Edman degradation sequencing. Once obtained, the oligonucleotide component (still full length 61-mer) of the cross-linked adduct needed to be substantially reduced. This is necessary in order to utilize the positive ion mode of the ESMS (oligonucleotides generate negative ions while peptides generate positive ions in ESMS), and to reduce potential noise from unwanted oligonucleotide fragmentation. To accomplish this, 60 pmol trypsin-digested 06.50/bFGF cross-linked adduct was digested with snake venom phosphodiesterase (SVP) and alkaline phosphatase (AP). SVP digests oligonucleotides into mononucleotides by sequentially cleaving the phosphodiester bond at the 5' and 3' ends of an oligonucleotide. AP removes PO_4^{3-} on the 3' and 5' ends of nucleotides by hydrolyzing them to hydroxyl groups. SVP/AP digestion of tryptic 06.50/bFGF should then yield a small nucleotide fragment (e.g. mono, di, or trinucleotide) cross-linked to a small peptide fragment (i.e., the fragment identified by Edman degradation).

Finally, 20 pmol of the SVP/AP/tryptic digested 06.50/bFGF cross-linked fragment was loaded onto an HPLC equipped ESMS/MS. ESMS showed a substantial peak at m/z 768.8 as shown in Figure R.5.5 on the following page. This peak was subsequently determined to be a doubly-charged ion due to the presence of the nearby peak of m/z 779.9. The peak at 779.9 is a result of the molecular ion associated with Na^+ in place of H^+. This alteration should result in an increase of m/z of 22. However, it appears as an increase in m/z of 11—half of that expected and thus the molecular ion must be doubly charged. As a result, this parent ion has a

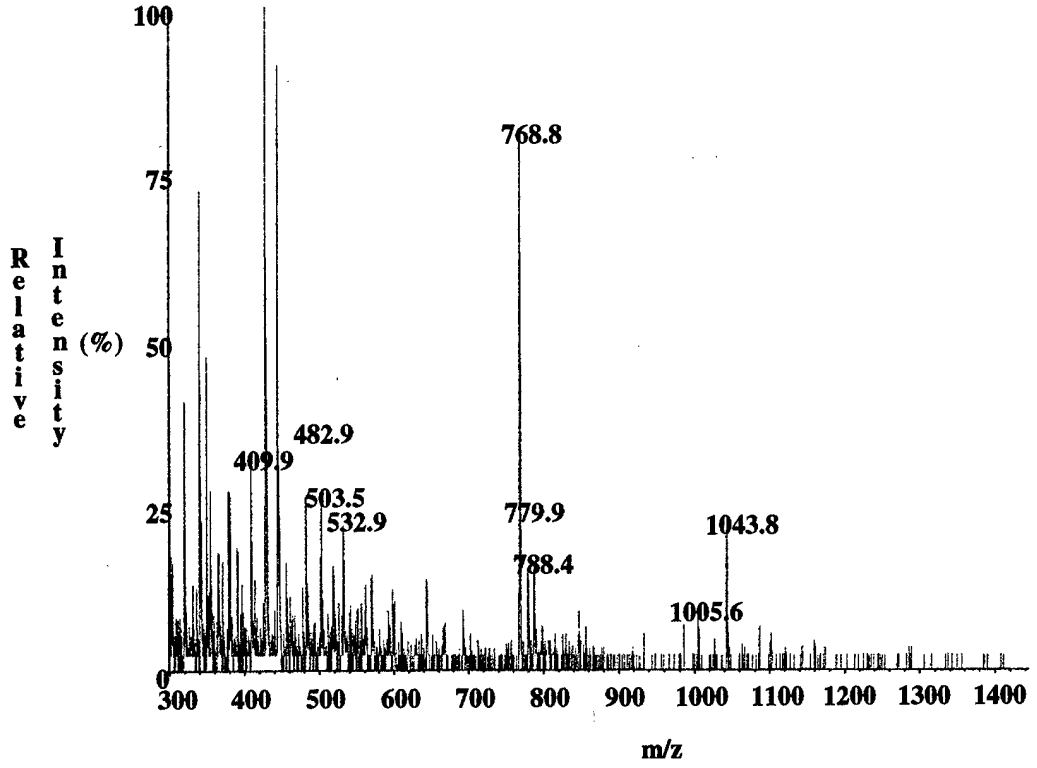


Figure R.5.5. ESMS of 06.50/bFGF SVP/AP/tryptic fragment. The adduct's molecular weight was determined to be 1537 (= 2x 768.8).

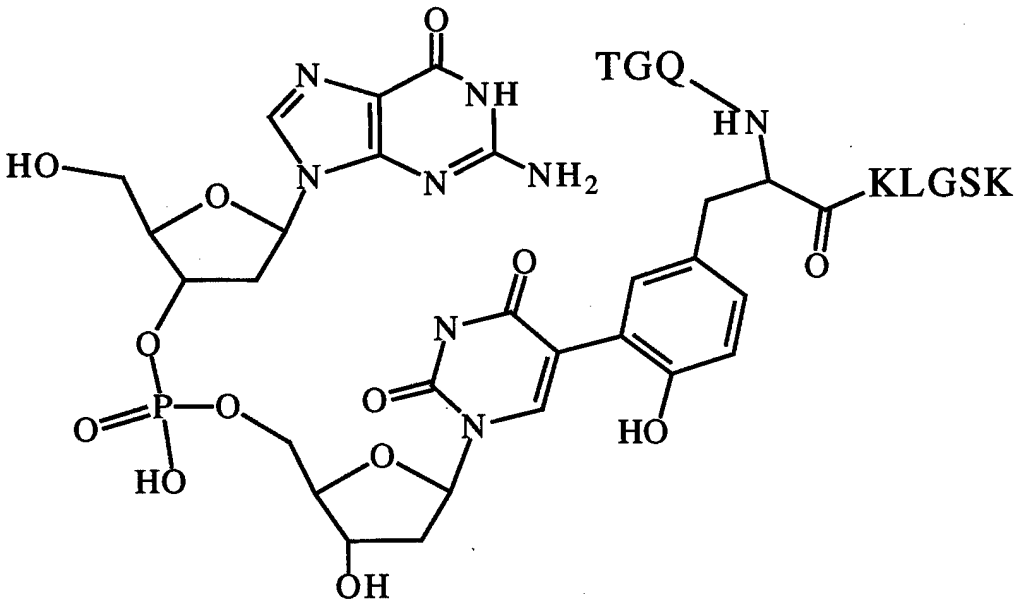


Figure R.5.6. GU (minus uracil's 5'-H) dinucleotide cross-linked to bFGF nonapeptide.

singly-charged m/z of 1537. This corresponds to a GU (minus uracil's 5'-H) dinucleotide covalently bound to Tyr within the nonapeptide fragment TGQN<u>Y</u>KLGSK whose structure is on the previous page in Figure R.5.6.

In order to confirm this as the structure of the molecular ion resulting from SVP/AP/trypsin digestion of the 06.50/bFGF cross-linked adduct, this ion was subjected to MS-MS. Figure R.5.7 below shows the resulting mass spectrum.

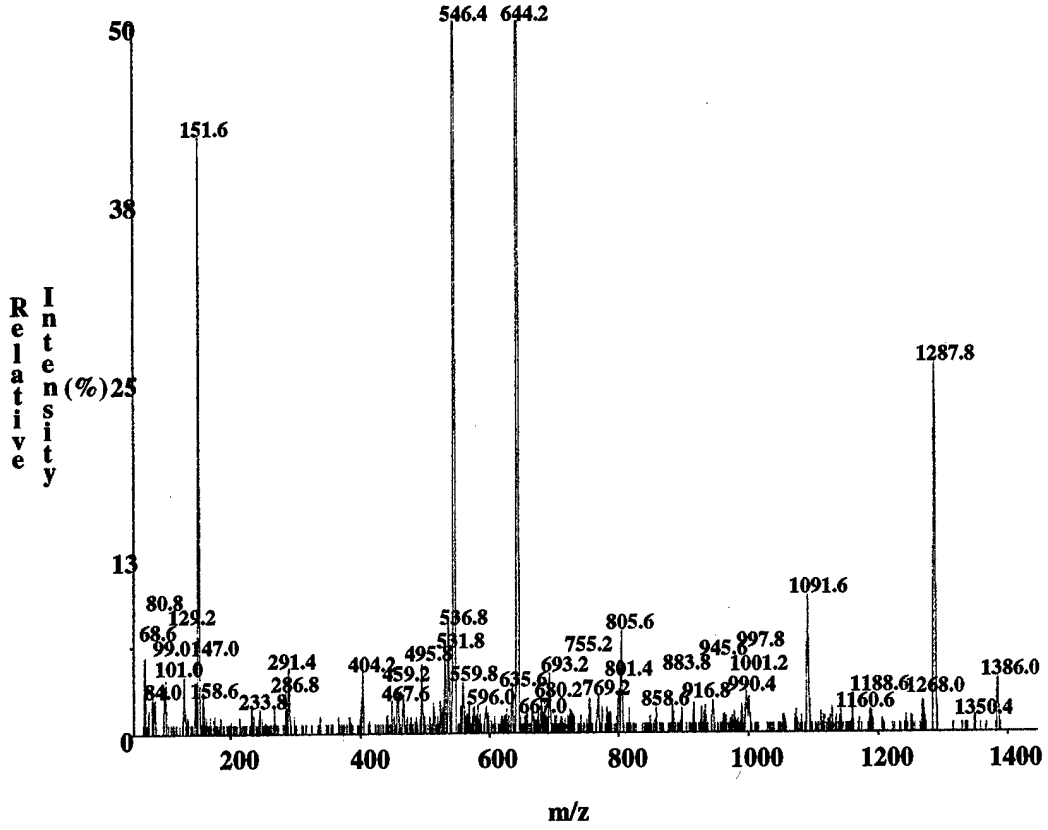
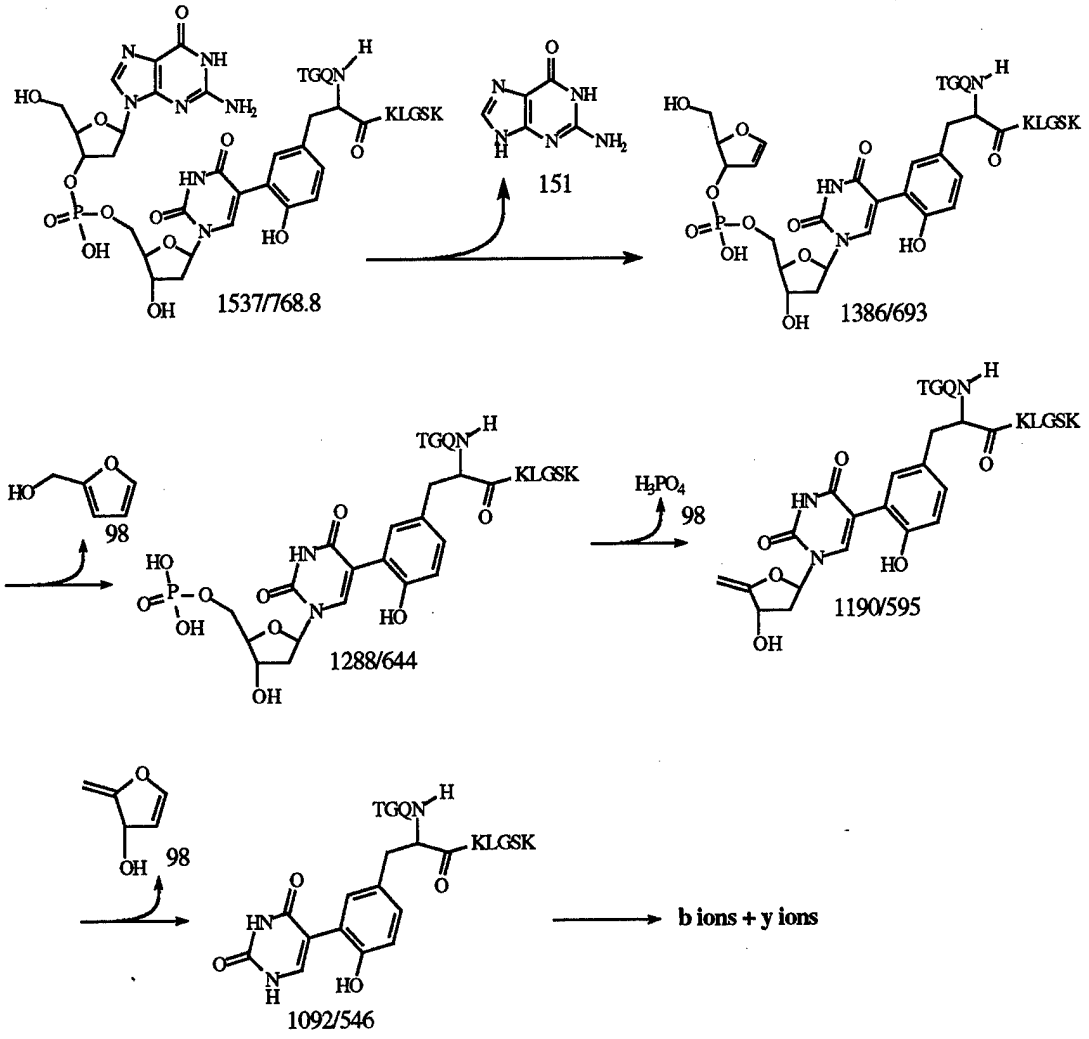


Figure R.5.7. MS-MS of molecular ion 768.8 m/z.

Finally, Scheme R.5.3 illustrates the daughter ions resulting from EI bombardment of this putative adduct. The corresponding m/z for these structures coincides exceedingly well with peaks in Figure R.5.7. As Scheme R.5.3 shows, fragmentation continues until the daughter ion with m/z 1092 (uracil base minus the 5'H covalently

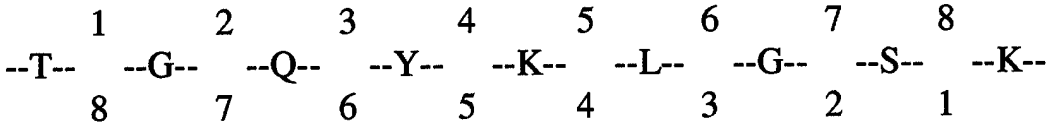


Scheme R.5.3. Mass spectrometric fragmentation of daughter ions from SVP/AP/trypsin digested 06.50/bFGF cross-linked adduct.

bound to the nonapeptide fragment) is generated. This daughter ion is sufficiently stable to permit generation of the ABC/XYZ ions resulting from peptide backbone bond cleavage with subsequent peptide sequencing possible. These results are shown in Table R.5.1 below. As can be seen here, the m/z of the the b and y-ions produced as N and C-terminal fragments correspond precisely to that expected from the non-

Table R.5.1. m/z of the b and y-ions resulting from EI fragmentation of 1092 m/z ion.

N-Terminal Fragments				*C-Terminal Fragments*			
n	*a*	*b*	*c*	*n*	*x*	*y*	*z*
0	---	1.01	16.02	9	---	1091.55	1074.52
1	74.06	102.06	117.07	8	1016.48	990.50	973.47
2	131.08	159.08	174.09	7	959.46	933.48	916.45
3	259.14	287.14	302.15	6	831.40	805.42	788.39
4	532.22	560.21	575.22	5	558.33	532.35	515.32
5	660.31	688.31	703.32	4	430.23	404.25	387.22
6	773.39	801.39	816.40	3	317.15	291.17	274.14
7	830.42	858.41	873.42	2	260.12	234.15	217.12
8	917.45	945.44	960.45	1	173.09	147.11	130.09
9	1045.54	1073.54	---	0	45.00	19.02	---



post-translationally modified nonapeptide bFGF fragment until the Tyr residue is reached. The mass of the cross-linked residue at this position is 110 m/z greater than the free y6 ion would be. This is precisely the mass of the uracil cross-linked to tyrosine minus 2H (one H from the 5' position of uracil which was originally bonded to BrdU and one H from a position ortho to the hydroxy on tyrosine's aromatic ring—the position of the new covalent bond). These results completely confirm Tyr133 as bFGF's cross-linking amino acid to the 06.50 aptamer.

The facility of the ESMS/MS technique to rapidly sequence a small quantity of photocross-linked oligonucleotide/protein adduct is an exciting first in peptide sequencing technology. These results reaffirm the usefulness of ESMS to sequence post translationally modified aptamers. More importantly, they enhance the usefulness of photocross-linking oligonucleotide/protein complexes as a tool of structure elucidation.

CHAPTER VIII: MAXAM-GILBERT DETERMINATION OF THE CROSS-LINKING POSITION OF 06.50

The final aspect to complete characterization of the covalently bound oligonucleotide/protein adduct is the determination of the cross-linking position on the oligonucleotide. This not only helps elucidate the oligonucleotide contact point and thereby important structural characteristics of the adduct, it can also be useful in any post-SELEX modification (e.g. replacing a BrdU with IdU at the cross-linking position in an attempt to increase cross-link yield while retaining robust photochemistry). In the past, cross-linking positions have been determined by single substitution experiments.

As described in the Introduction, oligonucleotide modification by incorporation of photocross-linking chromophores has helped identify DNA/protein contact points. In these experiments, a single photo-active base is incorporated into an affinity-binding oligonucleotide sequence for a given protein with subsequent irradiation and potential covalent bond formation. As such, there is no question regarding the oligonucleotide cross-linking position. Recently, several research groups investigating affinity SELEX have performed similar photocross-linking experiments to identify the oligonucleotide contact point for the protein target (Ruckman et al., 1998). These experiments involve the preparation of a series of modified versions of an affinity SELEX-evolved aptamer via solid phase synthesis. Each version will incorporate a single halouridine photocross-linking chromophore at a position originally occupied by a thymidine (DNA) or uridine (RNA) (other nucleotides may be modified as well depending on the chromophore of interest). Once prepared, each aptamer version will be irradiated with the protein at the

appropriate wavelength light with the hope that one version will yield a cross-link adduct. The procedure is illustrated in Figure R.6.1 below.

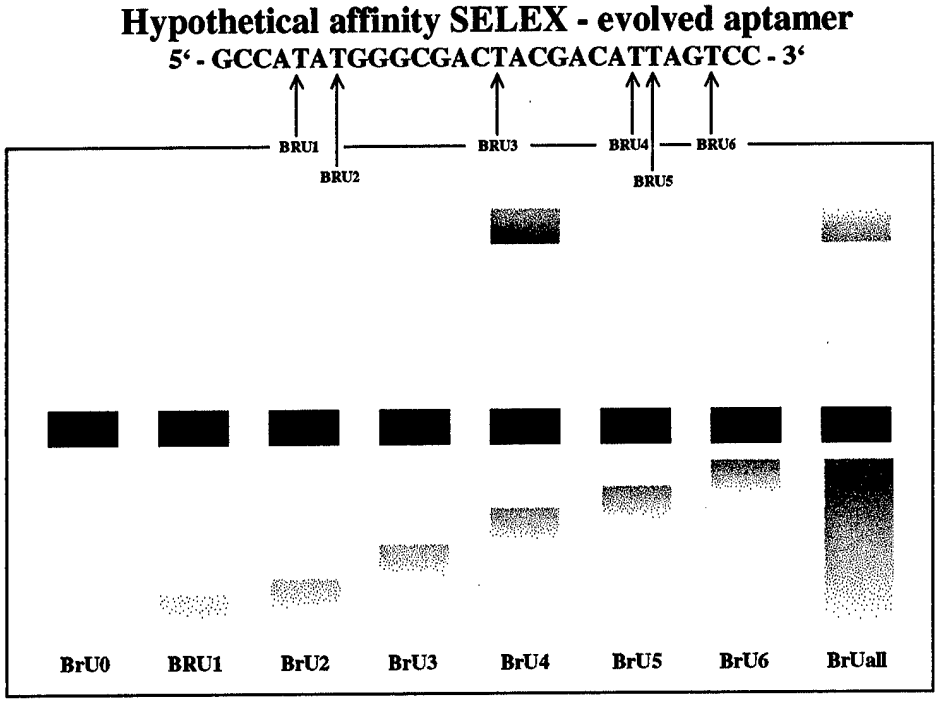


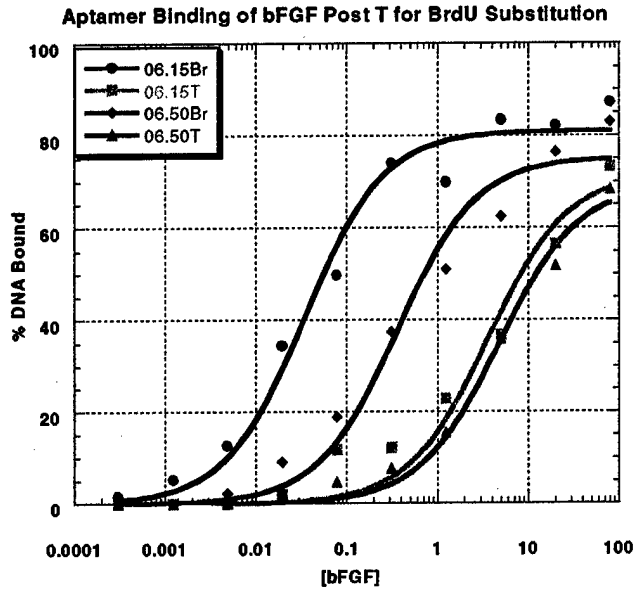
Figure R.6.1. The single substitution experiment to identify the cross-linking position of a post-affinity SELEX modifed aptamer.

This technique is precisely that used by Collins et al., (manuscript in preparation) in the identification of the cross-linking position of the aptamer 225-61. However, the ability of such post-affinity SELEX modifications to identify photocross-linking oligonucleotides or contact points in affinity SELEX-evolved aptamers is quite limited since cross-linking will only occur to certain amino acids in the target protein which possess a unique orientation with respect to the chromophore. Since this is not a characteristic specifically selected for in affinity SELEX, the ability to form a photo-covalent bond is sheer serendipity. Moreover, such post-affinity

SELEX modifications can perturb the oligonucleotide/protein interface with subsequent diminution of affinity binding and any cross-linking potential. Despite these limitations, single substitution modification has identified the oligonucleotide/protein contact points in numerous wild type systems and in a few affinity SELEX-evolved aptamers as well. However, it was not known whether post-PhotoSELEX modifications could be successfully employed to identify the oligonucleotide cross-linking position for identified aptamers.

The post-PhotoSELEX analog to single-substitution post-affinity SELEX modification is multiple substitution. That is, rather than replacing a single thymidine in an affinity-selected ssDNA aptamer with bromo-deoxyuridine, post-PhotoSELEX modification calls for the substitution of all BrdU positions with T except for one in each version of the aptamer. Several versions would then be synthesized until all original BrdU positions are singly represented. To assess the feasibility of such post-PhotoSELEX modifications, the impact of converting all BrdU positions to T on the binding affinity of the identified aptamers was measured.

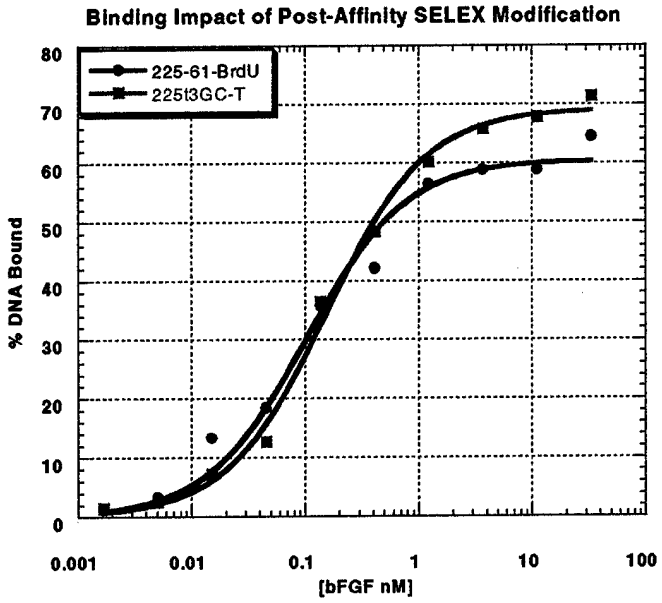
As mentioned above, photocross-linking ability can be severely impaired or even destroyed when binding affinity is reduced, so it was necessary to determine the impact of multiple BrdU→T substitions on the binding affinity of 06.15 and 06.50 for bFGF. This was done via PCR amplification of both sequences using standard dNTPs (i.e. TTP was included and BrdUTP was removed). The resulting binding affinity of these aptamers was measured for bFGF and is shown in Figure R.6.2 on the following page. In both cases, an order of magnitude reduction in binding affinity was observed. This is additional confirmation as to the exquisite ability of PhotoSELEX to specifically identify high affinity binding/high quantum yield cross-



Aptamer	06.15Br	06.15T	06.50Br	06.50T
K_d	0.016 nM	0.178 nM	0.56 nM	6.31 nM

Figure R.6.2. Impact of multiple BrdU→T substitutions on binding affinity for 06.15 and 06.50 to bFGF.

linking aptamers for a protein target. Clearly, these aptamers benefit from the electonegativity of Br (increased dipole-dipole bonding potential) in forming affinity-based associations with bFGF; they were selected, in part, for this characteristic. Their replacement with a similar sized but relatively electropositive methyl group destroys the dipole-dipole bonds so important to the exceptional affinity enjoyed by these aptamers. Interestingly, at least in the case of 225-61, the post-affinity SELEX modification of T→BrdU does not impact binding as is shown in Figure R.6.3 on the following page. In this case, the wild type aptamer evolved without the advantage of



Aptamer	225t3GC-T	225-61-BrdU
K_d	0.25 nM	0.195 nM

Figure R.6.3. Impact of multiple T→BrdU substitutions on binding affinity of 225-61 to bFGF (225t3GC-T is a truncated version of 225-61 identified by Collins and coworkers (manuscript in preparation).)

the electronegative Br, and as such, does not utilize its dipole-dipole bond-forming ability in achieving affinity. As a result, the T→BrdU modification in this case is a transparent one.

While all this helps bolster the claim that PhotoSELEX rivals if not surpasses affinity SELEX in all aspects of aptamer identification, the problem of identifying the oligonucleotide cross-link position remained. Fortunately, it seemed possible to identify the cross-link position if the protein could be substantially (but not

completely) digested enzymatically followed by a systematic cleavage of the oligonucleotide itself. The Maxam-Gilbert sequencing technique provides just this kind of systematic cleavage.

Maxam-Gilbert sequencing was invented by Allan Maxam and Walter Gilbert in 1980 to sequence DNA. Its objective is to chemically cleave 5'-^{32}P-labelled DNA in a base-specific manner, so that, on average, each DNA strand is randomly cleaved at only one susceptible position. This yields a series of DNA fragments which correspond in length to the position of a particular base from the 5' end. For instance, the Maxam-Gilbert sequencing of the oligonucleotide:

^{32}P-GAGCTTCAGGCCGTCAT-3'

under conditions which specifically cleave at the 5' end of cytidine nucleotides would result in the following DNA fragments:

^{32}P-GAGCTTCAGGCCGTCAT-3'

^{32}P-GAGCTTCAGGCCGT

^{32}P-GAGCTTCAGGC

^{32}P-GAGCTTCAGG

^{32}P-GAGCTT

^{32}P-GAG

These fragments can then be separated by PAGE and the C-positions read directly from an autoradiogram. A full complement of such base-specific cleavage reactions permits determination of the entire oligonucleotide sequence. An oligonucleotide cross-linked to bFGF with subsequent digestion of the protein would be chemically cleaved in an identical manner. The only difference in appearance between the Maxam-Gilbert sequencing of the noncross-linked and cross-linked versions of a

given aptamer would be a gel shift at the cross-linking base corresponding in magnitude to the size of the protein fragment. Since a moderate but significant shift would permit ready identification of the cross-link position, it seemed a peptide fragment approximately ten amino acids long would be ideal. Conveniently, this is just about the peptide fragment remaining after trypsin digestion of the 06.50/bFGF cross-linked adduct. After the shift, the Maxam-Gilbert sequencing gel for the digested cross-linked adduct would again appear identical to the noncross-linked version except all bands would be shifted up the gel a constant amount corresponding again to the peptide fragment. Figure R.6.4 shown below illustrates the conceptualized Maxam-Gilbert sequencing gel of both the noncross-linked and cross-linked digested versions of a hypothetical cross-linking aptamer.

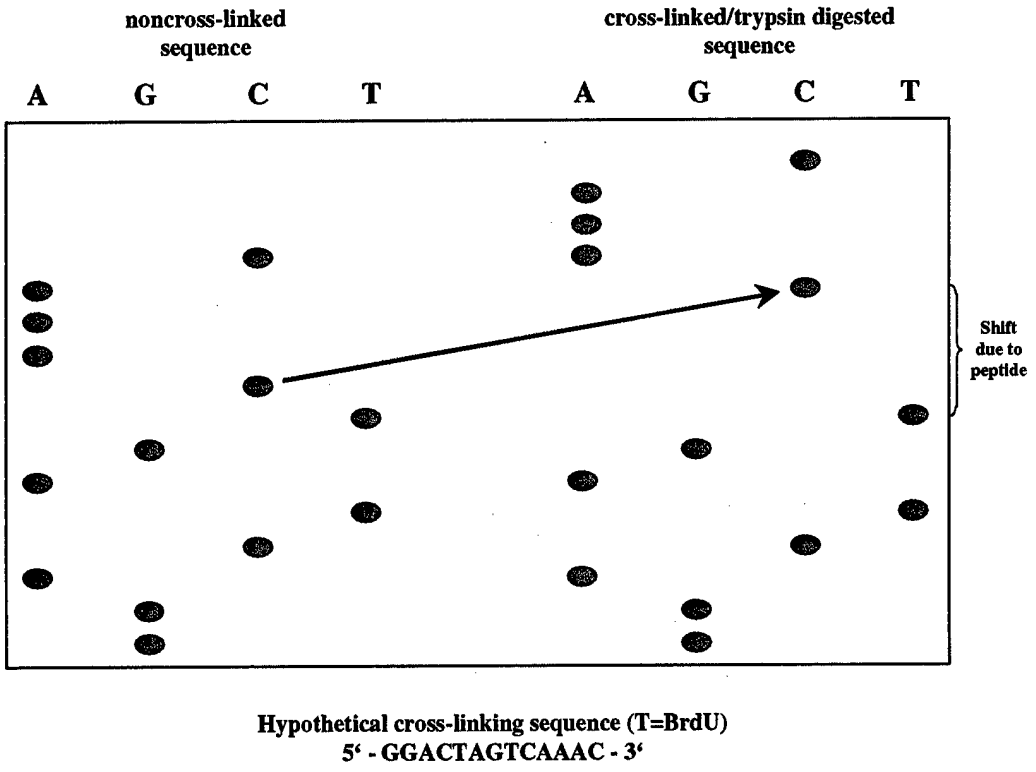
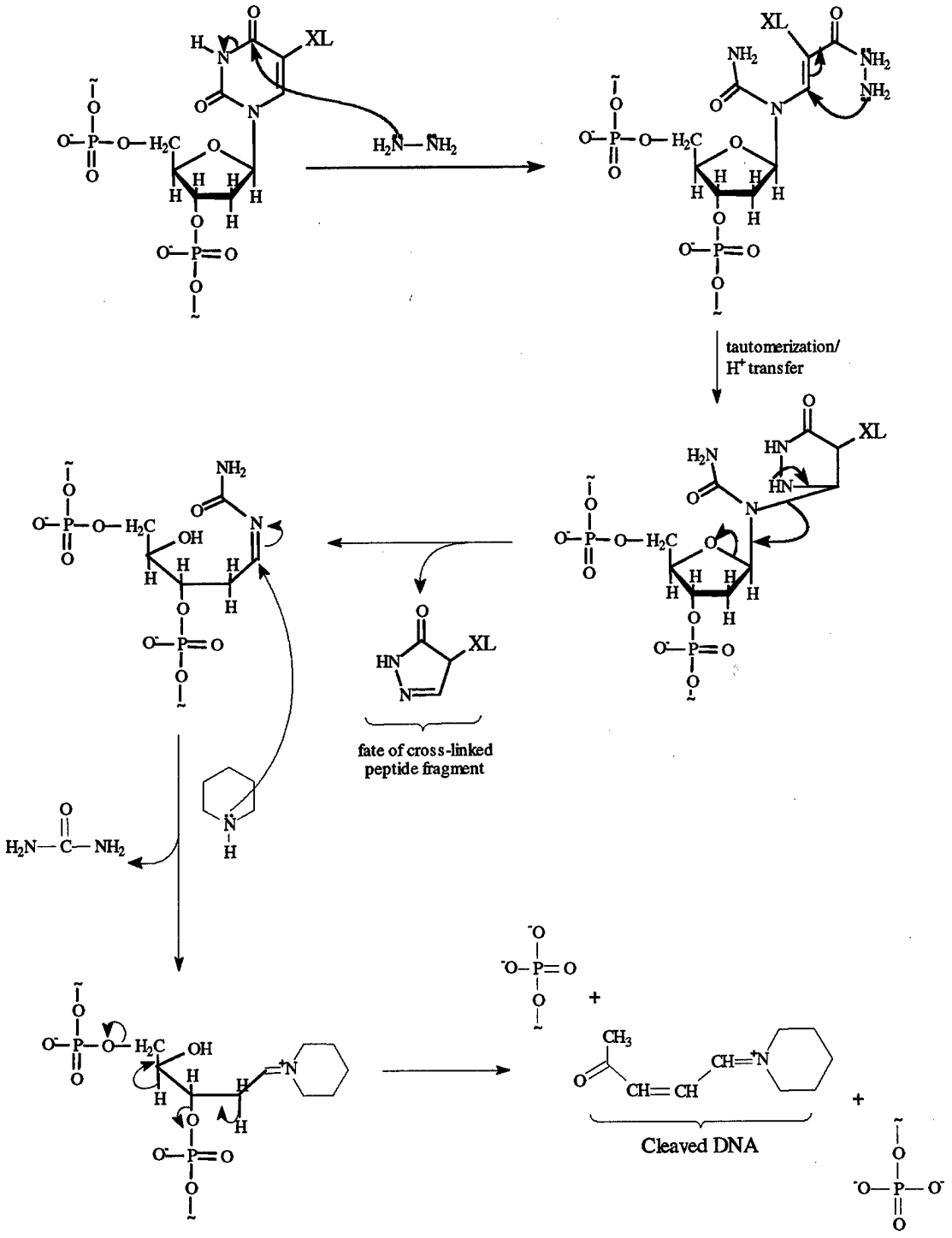


Figure R.6.4. Maxam-Gilbert sequencing determination of the cross-link position on a hypothetical PhotoSELEX-evolved aptamer.

Clearly, the cross-link to bFGF must be at one of the BrdU positions in the sequence. However, it is not immediately clear whether the cross-linking



Scheme R.6.1. Mechanism of hydrazine ring opening of thymidine and subsequent DNA phosphate backbone cleavage by piperidine (after Voet and Voet, 1995).

chromophore would appear at the bottom as (shown in Figure R.6.4) or at the top of the peptide shift region. An ability to predict the appearance of the peptide shift region would enhance confidence in the reliability of Maxam-Gilbert sequencing as a method to identify the cross-linking position in PhotoSELEX-evolved aptamers.

The appearance of the shift region can be predicted based upon the mechanism of chemical cleavage of thymidine nucleotides by hydrazine and piperidine. As can be seen in Scheme R.6.1 on the previous page, the tryptic fragment which is covalently bound to the cross-linking nucleotide is removed prior to scission of the phosphate backbone. Therefore, the cross-linking nucleotide in the aptamer should appear at the bottom of the peptide shift region.

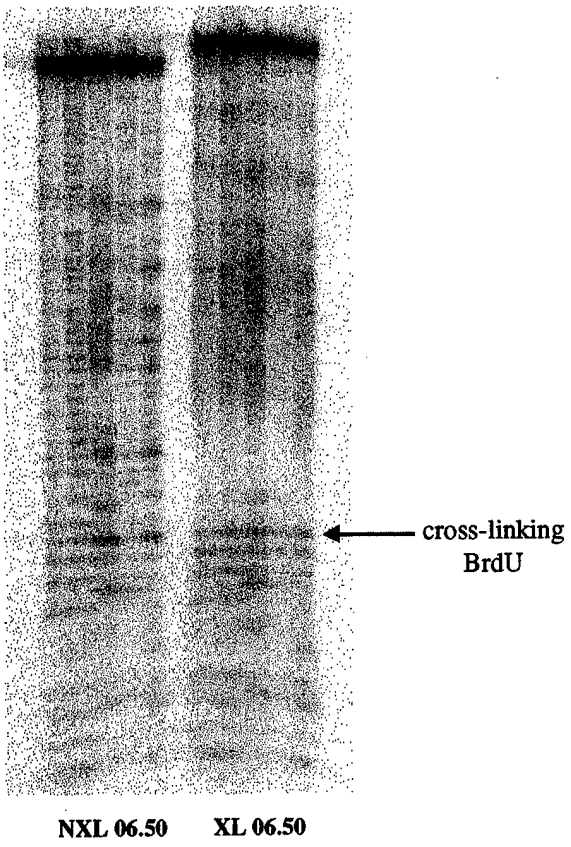


Figure R.6.5. Maxam-Gilbert sequencing determination of the cross-link position of 06.50.

The Maxam-Gilbert sequencing experiment was conducted with noncross-linked and cross-linked 06.50; the PAGE partition is shown on the preceding page in Figure R.6.5. The PAGE partition clearly shows a peptide shift region for cross-linked 06.50. In addition, the shift corresponds to a BrdU position and the putative cross-linking nucleotide is at the bottom of the shift region in agreement with the prediction from Scheme R.6.1. The ability of Maxam-Gilbert DNA sequencing to identify the nucleotide cross-linking position of 06.50 (shown in Figure R.6.6 below) was yet another extremely positive development for the PhotoSELEX technology.

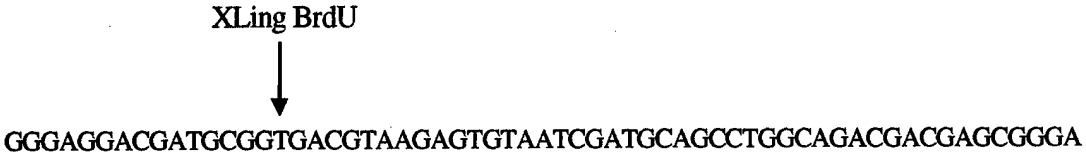


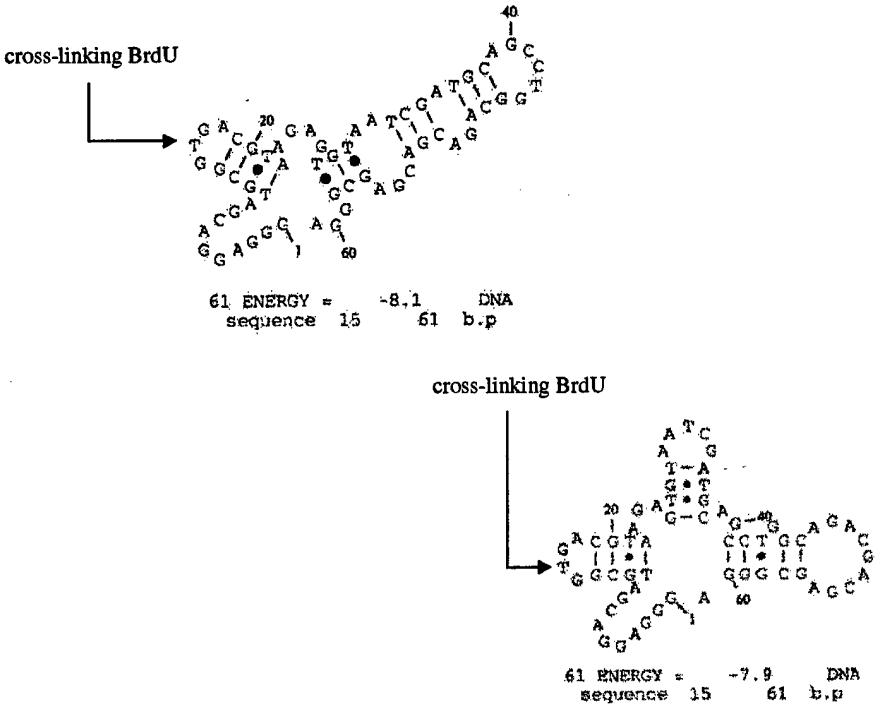
Figure R.6.6. Illustration of the 06.50 cross-linking position within the aptamer sequence (T = BrdU in random region).

Rather than expending considerable time and resources as well as creating substantial amounts of hazardous waste to generate singly-substituted versions of a given aptamer, Maxam-Gilbert sequencing provides cross-linking nucleotide identification in a single experiment. Moreover, an interesting and unexpected consequence of this procedure appears in the PAGE lane corresponding to cleavage at A>C residues in Figure R.6.5.

A cursory examination of this lane shows that seven bands appear much more prominently than any other bands in that lane. These bands are in fact cleavage reactions at the BrdU locations in 06.50. Evidently, the electronegative Br makes the

uracil base much more reactive than thymine, adenine or cytosine. As mentioned earlier, since the cross-linking position must occur at one of the BrdU positions, it seems it would be possible to complete the Maxam-Gilbert sequencing and subsequent identification of the cross-linking position of an oligonucleotide using a single base-specific cleavage reaction.

Finally, the identification of the cross-linking nucleotide in 06.50 appears to corroborate the structures generated by the MulFold program reported earlier. Recall from the first section of this chapter that this base position was placed in protein-accessible hairpin loops for the two most stable secondary structures for 06.50 which are shown here again.



R.2.1. MulFold secondary structures for 06.50.

CHAPTER IX: CONCLUSIONS

The utility of the combinatorial chemistry technique known as the Photochemical Systematic Evolution of Ligands by Exponential Enrichment was investigated as a potential identifier of therapeutic and diagnostic aptamers. The results presented in this dissertation indicate great promise for this nascent technology. In particular, the results show:

1. PhotoSELEX can be accomplished against a target for which no advantage derived from biased randomization or prior affinity selection is available.

2. PhotoSELEX possesses an exquisite ability to identify the best photocross-linking aptamers from among all aptamers which cross-link a given target.

3. PhotoSELEX can select for affinity binding while selecting for photocross-linking ability.

4. PhotoSELEX can be performed in a systematic, reproducible manner without *in situ* modifications.

5. PhotoSELEX can identify high affinity binding aptamers in fewer rounds than can affinity SELEX.

In addition, a statistical calculation indicated that there is a high probability (69%) of generating more than one copy of a given sequence during synthesis of a 30-mer randomized oligonucleotide library. As a result, identification of the same aptamer during separate PhotoSELEX experiments derived from such a library for a given target may reflect the extraordinary capability of PhotoSELEX to sample sequence space and rapidly evolve the highest quantum yield cross-linking/highest affinity binding aptamers.

This research also sought to develop new techniques for characterizing and evaluating aptamers identified by PhotoSELEX. First, SDS-PAGE was shown to be an effective method for eliminating bFGF's propensity to form multimer aggregates

and thereby assess an aptamer's true cross-link yield. For example, the cross-link yield for aptamer 06.50 with bFGF at protein excess was 72%--among the highest ever recorded for a BrdU→T ssDNA/protein system.

Second, the diagnostic potential of two unique aptamers evolved by PhotoSELEX for bFGF was demonstrated. These aptamers were shown to give measurable cross-link yield at exceptionally low [bFGF] (0.0032nM) when measured by SDS-PAGE; cross-link yield was even detectable at this concentration in the presence of diluted serum. They were also shown to readily distinguish among consanguineous proteins based upon affinity. In addition, a diagnostic apparatus was postulated which would take advantage of the remarkable capabilities exhibited by these aptamers.

Third, the protein's cross-linking amino acid was identified by electrospray mass spectrometry. Through enzymatic digestion of both the protein and the oligonucleotide components of a cross-linked adduct, a fragment suitable for mass spectrometric analysis was generated. The molecular weight of this intact ion was subsequently determined by electrospray mass spectrometry, and the peptide sequenced by associated MS/MS. The daughter ions generated by MS/MS then revealed the cross-linking peptide to be Tyr133--an amino acid in the heart of bFGF's heparin binding site. The mass spectrometric sequence results were completely confirmed by Edman degradation sequencing.

Finally, an aptamer's cross-linking chromophore position was identified by Maxam-Gilbert sequencing. After tryptic digestion of the protein component, the cross-linked adduct was subjected to the base-specific phosphate backbone cleavage reactions of Maxam-Gilbert sequencing. The cross-linking nucleotide appeared at the

bottom of a peptide fragment shift region within the oligonucleotide's sequence and was thereby readily identified. In the case of 06.50, the cross-linking nucleotide is the first BrdU position in the random region. Confidence in the validity of these results is enhanced due to agreement with MulFold structures of 06.50 which place this nucleotide in protein-accessible hairpin loops. Moreover, the Maxam-Gilbert reaction conditions for A>C cleavage yielded unusually substantial cleavage at all BrdU positions in the 06.50 aptamer. It was therefore postulated that subsequent PhotoSELEX-evolved aptamers could be characterized by this reaction alone (since cross-linking must occur at a chromophore). This technique might also be used as a convenient method to rapidly assess the degree of convergence in a PhotoSELEX experiment.

As the reader might readily deduce from the chapter on the PhotoSELEX experiments, a great deal of time and effort was expended elucidating all of the factors necessary to permit PhotoSELEX convergence (e.g. PCR conditions, maximal protein digestion, successful PAGE partition, etc.). When one considers that any one of these factors was capable of rendering PhotoSELEX untenable, it is remarkable that the experiment worked at all. Now that the results have been compiled and the PhotoSELEX employing the 5-bromo-2'-deoxyuridine chromophore for the bFGF target has succeeded beyond its originators' dreams, it seems nothing short of a miracle. Unfortunately, its extraordinary ability to rapidly evolve high quantum yield/high affinity binding aptamers for a given target may still be its downfall.

Several researchers in the biotech industry had tried their hand for several months at PhotoSELEX and failed to see any degree of evolution. The reason for their failure is now clear. An inability or unwillingness to *systematically* explore and

employ the PhotoSELEX protocol virtually guarantees its downfall due to any one of the crucial factors mentioned above. In short, PhotoSELEX is exceptionally fastidious which does not necessarily suit it for use in the "results now" mentality of industry. Fortunately, there does appear a way to utilize the amazing sampling and amplifying power of PhotoSELEX while mitigating its capricious temperament. The next step forward in developing SELEX as a combinatorial chemistry technique should be a careful examination of ThermalSELEX--SELEX based upon the formation of a thermal covalent bond.

BIBLIOGRAPHY

Abraham, J.A., Whang, J.L., Tumolo, A., Mergia, A., Friedman, J., Gospodarowicz, D. and Fiddes, J.C. (1986). *Embo. J.*, **5**: 2523-2528.

Allen, T., Wick, K. and Matthews, K. (1991). *J. Biol. Chem.*, **266**: 6113-6119.

Amouyal. E., Bernas, A. and Grand, D. (1979). *Photochem. Photobiol,* **29**: 1071-1077.

Baird, A. and Klagsbrun, M. (eds.) (1991). The fibroblast growth factor family. Vol. 638, New York Academy of Sciences, New York.

Bennett, S.E., Jensen, O.N., Barofsky, D.F. and Mosbaugh, D.W. (1994). *J. Biol. Chem.*, **269**: 21870- 21879.

Biemann, K. (1988). *Mass Spectrom.*, **16**: 99.

Burgess, W.H. and Maciug, T. (1989). *Annu. Rev. Biochem.*, **58**: 575-606.

Burgess, W.H., Shaheen, A.M., Ravera, M., Jaye, M., Donohue, P.J. and Winkles, J.A. (1990). *J. Cell. Biol.*, **111**: 2129-2138.

Collins, B.D., Breines, K.M., Jellinek, D., Hicke, B.J., Ciftan, S.A., Janjic, N., Thorp, H.H. and Willis, M.C. (Manuscript in preparation).

Connor, D.A., Fallick, A.M. and Shetlar, M.D. (1998a). *Photochem. Photobiol.*, **68**: 1-8.

Connor, D.A., Fallick, A.M., Young, M.C. and Shetlar, M.D. (1998b). *Photochem. Photobiol.*, **68**: 299-308.

Cook, K.S., Fisk, G.J., Hauber, J., Usman, N., Daly, T.J. and Rushe, J.R. (1991). *Nucleic Acids Res,*. **19**: 1577-1583.

Creed, D. (1984). *Photochem. Photobiol.*, **39**: 563-575.

Dietz, T.M. and Koch, T.H. (1987). *Photochem. Photobiol.*, **46**: 971-978.

Dietz, T.M., Von Tebra, R.J., Swanson, B.J. and Koch, T.H. (1987). J. Am. Chem. Soc., **109**: 1793-1797.

Dirix, L.Y., Vermeulen, P.B., Pawinski, A., Prove, A., Benoy, I., Depouter, C., Martin, M. and Vanoosterom, A.T. (1997). *Brit. J. Cancer,* **76**: 238-243.

Ellington, A.D. and Szostak, J.W. (1990). *Nature*, **346**: 818-822.

Engvall, E., et. al. (1971) Immunochemistry, **35**: 350-356.

Eriksson, A.E., Cousens, L.S. and Mathews, B.W. (1993). *Protein Sci.*, **2**: 1274-1284.

Eriksson, A.E., Cousens, L.S., Weaver, L.H. and Mathews, B.W. (1991). *Proc. Natl. Acad. Sci.*, USA. **88**: 3441-3445.

Esch, R., Baird, A., Ling, N., Veno, H., Hill, F., Deneroy, L., Klepper, R., Gospodarowicz, D., Boehlen, P. and Buillemin, R. (1985). *Proc. Natl. Acad. Sci.*, USA. **82**: 6507-6511.

Faham, S., Hileman, R.E., Fromm, J.R., Linhardt, R.J. and Rees, D.C. (1996). *Science,* **271**: 1116-1120.

Famulok, M., and Szostak, J. (1992). *J. Am. Chem. Soc.*, **114**: 3990-3991.

Fitzwater, T. and Polisky, B. (1996). *Meth. in Enzym.*, **267**: 275-301.

Folkman, J. (1970). Carcinoma of the colon and antecedent epithelium. Burdette, W.J. (ed.), pp. 113-127. Springfield, IL.

Folkman, J. (1971). *N. Eng. J. Med.*, **285**: 1182.

Folkman, J. and Klagsbrun, M. (1987). *Science,* **235**: 442-447.

Folkman, J., Long, D.M. and Becker, F.F. (1963). *Cancer*, **16**: 453.

Franklin, J.L., Dillard, J.G., Rosenstock, H.M., Herron, J.T., Draxl, K. and Field, F.H. (1969). *NSRDS-NBS*, **26**: 129.

Gimbrone, M.A. (1969). *Nature*, **222**: 33.

Green, L.S., Jellinek, D., Jenison, R., Ostman, A., Heldin, C.H. and Janjic, N. (1996). *Biochemistry*, **35**: 14413-14424.

Greenblatt, M. and Shubik, P. (1968). *J. Natl. Cancer Inst.*, **41**: 111.

Gold, L. (1995). *J. Biol. Chem.*, **270**: 13581-13584.

Gold, L., Polisky, B., Uhlenbeck, O. and Yarus, M. (1995). *Annu. Rev. Biochem.*, **64**: 763-797.

Gospodarowicz, D. (1991). *Ann. N.Y. Acad. Sci.*, **638**: 1-8.

Gospodarowicz, D., Bialecki, H. and Greenburg, G. (1978). *J. Biol. Chem.*, **253**: 8736.

Gospodarowicz, D. and Cheng, J. (1986). *J. Cell. Physiol.*, **128**: 475-484.

Gott, J.M., Willis, M.C., Koch, T.H. and Uhlenbeck, O.C. (1991). *Biochemistry*, **30**: 6290-6295.

Harper, J.W., and Lobb, R.R. (1998). *Biochemistry*, **27**: 671-678.

Harrison, C.A., Turner, D.H. and Hinkle, D.C. (1982). *Nucleic Acids Res.*, **10**: 2399-2415.

Heaphy, S., Dingwalk, C., Ernverg, I., Gait, M. and Skinner, M.A. (1990). *Cell*, **60**: 685-693.

Hertig, A.T. (1935). *Contrib. Embryology*, **25**: 37.

Hicke, B.J., Willis, M.C., Koch, T.H. and Cech, T.R. (1994). *Biochemistry*, **33**: 3364-3373.

Hsu, P.I. (1997). *Anticancer Res.*, **17**: 2803-2809.

Irvine, D., Tuerk, C. and Gold, L. (1991). *J. Mol. Biol.*, **222**: 739.

Ito, S., Sarto, I. and Matsuura, T. (1980). *J. Am. Chem. Soc.*, **102**: 7535-7541.

Jaeger, J.A., Turner, D.H. and Zucker, M. (1989). Proc. Natl. Acad. Sci., USA. *Biochem.*, **86**: 7706-7710.

Jaeger, J.A., Turner, D.H. and Zucker, M. (1989). *Meth. Enzymal.*, **183**: 281-306.

Jellinek, D., Lynott, C.K., Rifkin, D.B. and Janje, N. (1993). *Proc. Natl. Acad. Sci.*, **90**: 11227-11231.

Jenison, R.D., Gill, S.C., Pardi, A. and Polisky, B. (1994). *Science*, **24**: 234-241.

Jensen, K.B., Green, L., MacDougal-Waugh, S. and Turek, C. (1994). *J. Mol. Biol.*, **235**: 237-247.

Jensen, K.B., Atkinson, B.L., Willis, M.C., Koch, T.H. and Gold, L. (1995). *Proc. Natl. Acad. Sci.*, **92**: 12220-12224.

Jensen, O.N., Barofsky, D.F., Young, M.C., Vonltippel, P.H., Swenson, S. and Seifried, S.E. (1993). *Rapid Commun. Mass Spectrom.*, **7**: 496-501.

Jensen, O.N., Barofsky, D.F., Young, M.C., Vonltippel, P.H., Swenson, S. and Seifried, S.E. (1994). Techniques in Protein Chemistry V. pp. 27-37, Academic Press, New York.

Jensen, O.N., Kulkarni, S., Aldrich, J.V. and Barafsky, D.F. (1996). *Nucleic Acids Res.*, **24**: 3866-3872.

Jovanovic, S.V., Harriman, A. and Simil, M.G. (1986). *J. Phys. Chem.*, **90**: 1935-1939.

Judson, H.F. (1979). The eigth day of creation. New York: Simon & Schuster.

Karn, J., Dingwall, C., Finch, J.T., Heaphy, S. and Gait, M.J. (1991). *Biochemistry*, **73**: 9-16.

Kastrup, J.S., Eriksson, E.S., Dalboge, H. and Flodgaard, H. (1997). *Acta Cryst.*, **D53**: 160-168.

Katousian-Safadi, M., Laine, B., Chartier, F., Cremet, J., Belaiche, D., Sautiere, P. and Charlier, M. (1991). *Nucleic Acids Res.*, **19**: 4937-4941.

Katousian-Safadi, M. and Charlier, M. (1994). *Biochimie*, **76**: 129-132.

Khalili, K., Rappaport, J. and Khoury, G. (1988). *Embo. J.*, **7**: 1205-1210.

Kim. S.H., Suddath, F.L., Quisley, G.J., McPherson, A. and Sussman, J.L. (1974). *Science,* **185**: 435-440.

Kin, M., Sata, M., Veno, T. and Torimura, T. (1997). *J. of Hepatology.*, **27**: 677-687.

Kjellen, L. and Lindahl, U. (1991). *Ann. Rev. Biochem.*, **60**: 443-475.

Klagsbrun, M., Knighton, D. and Folkman, J. (1976). *Cancer Res.*, **36**: 110

Kuchler, K. (1993). *TICB,* **3**: 3-421.

Markowitz, A. (1972). *Biophys. Acta.*, **281**: 522-534.

Matsuo, T., Mihara, S. and Veda, I. (1976). *Tetrahedron Lett.*, p. 2581.

Maxam, A.M. and Gilbert, W. (1980). *Methods Enzym.,* **65**: 499-650.

McLafferty, F.W., Bente, P.F., Kornfeld, R., Tsai, S.C. and Howe, I (1973a). *J. Am. Chem. Soc.*, **95**: 2120.

McLafferty, F.W., Bente, P.F., Kornfeld, R., Haddon, W.F., Levsen, K., Sakai, I., Benete, P.F., Tsai, S.C. and Schuddemage, H.D.R. (1973b). *J. Am. Chem. Soc.*, **95**: 3886.

Meisenheimer, K.M. and Koch, T.H. (1997). *Cit. Rev. Biochem. and Mol. Biol.*, **32**: 101-140.

Merrill, B.M., Stone, K.L., Cobianchi, F., Wilson, S.H. and Williams, K.R. (1988). *J. Biol. Chem.*, **263**: 3307-3313.

Mignatti, P. et al. (1991). *Proc. Natl. Acad. Sci., USA*, **88**: 11007.

Mignatti, P. et al. (1992). *J. Cell Physiol.*, **151**: 81.

Moscatelli, D. (1988). *J. Cell Biol.*, **107**: 753-759.

Moy, F.J., Safran, M., Seddon, A.P., Kitchen, D., Bohlen, P., Aviezer, D., Yayon, A. and Powers, R. (1997). *Biochemistry*, **36**: 4782-4791.

Moy, F.J., Seddon, A.P., Bohlen, P. and Powers, R. (1996). *Biochemistry*, **35**: 13552-13561.

Nguyen, M., Watanabe, H., Budson, A.E., Richie, J.P., Hayes, D.F. and Folkman, J. (1994). *Natl. Cancer Inst.*, **86**: 356-361.

Norris, C.L., Meisenheimer, K.M. and Koch, T.H. (1997). *Photochem. Photobiol.*, **65**: 201-207.

Ogata, R. and Gilbert, W. (1977). *Proc. Natl. Acad. Sci.*, USA, **74**: 4973-4976.

Osborne, S.E. and Ellington, A.D. (1997). *Chem. Rev.*, **97**: 349-370.

Paradiso, P.R., Nakashima, Y. and Konigsberg, W. (1979). *J. Biol. Chem.*, **254**: 4739-4744.

Prestrelski, S.J., Fox, G.M. and Arakawa, T. (1992). *Archives of Biochemistry and Biophysics*, **293**: 314-319.

Ramsey, B.G. (1979). *J. Organic Chem.*, **44**: 2093-2097.

Relf, M., Lejuene, S., Scott, P.A.E., Fox, S., Smith, K., Leek, R., Moghaddam, A., Whitchoun, R., Bicknell, R. and Harris, A.L. (1997). *Cancer Res.*, **57**: 963-969.

Roesptoff, P. and Fohlman, J. (1984). *Biomed. Mass Spectrom.*, **11**: 601.

Rothman, W. and Kearns, D.R. (1967). *Photochem. Photobiol.*, **6**: 775-778.

Ruckman, J., Green, S., Beeson, J., Waugh, S., Gillette, W.L., Henninger, D.D., Claeson-Welsh, L., Janjic, N. (1998). *J. Biol. Chem.*, **273**: 20556-10567.

Saito, I., Matsuura, T. and Helene, C. (1981). *Photochem. Photobiol.*, **33**: 15-19.

Saito, I. and Matsuura, T. (1985). *Acc. Chem. Res.*, **18**: 134-141.

Saksela, O., Moscatelli, D., Sommer, A. and Rifkin, O.E. (1998). *J. Cell. Biol.*, **107**: 743-751.

Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989). Molecular cloning: A laboratory manual (2nd ed.) Cold Spring Harbor Laboratory Press. Ford, N., Nalam, C., Ferguson, M. (eds.), Plainview, NY.

Schneider, D.J., Feigon, J. and Hostomsky, Z. (1995). *Biochemistry*, **34**: 9599-9610.

Shaw, A.A., Falick, A.M. and Shetlar, M.D. (1992). *Biochemistry*, **31**: 10976-10983.

Shing, Y., Folkman, J., Sullivan, R., Butterfield, C., Murray, J. and Klogsbrun, M. (1984). *Science*, **223**: 1296-1299.

Smith, K.C. (1964). *Photochem. Photobiol.*, **3**: 415-427.

Somei, M. and Natsume, M. (1973). *Tetrahedron Lett.*, p. 2451.

Sundberg, R.J. (1970). The chemistry of indoles. Academic Press. New York.

Swanson, B.J., Kutzer, J.C. and Koch, T.H. (1981). *J. Am. Chem. Soc.*, **103**: 1274-1276.

Tanner, N.K., Hanna, M.M. and Avelson, J. (1988). *Biochemistry*, **27**: 8852-8861.

Thompson, S.A. and Fiddes. (1991). *Ann. N.Y. Acad. Sci.*, **638**: 78-88.

Thompson, L.D., Pantoliano, M.W. and Springer, D.A. (1994). *Biochemistry*, **33**: 3831-3840.

Tuerk, C. and Gold, L. (1990). *Science*, **249**: 505-510.

Tuerk, C. MacDougal, S. Hertz, G. and Gold L. (1993). The polymerase chain reaction. Eds Ferre, R. Mullis, K., Gibbs, R. and Ross, A. Birkhauser, Basel and Springer, New York.

Tuerk, C. and MacDougal-Waugh, S. (1993). *Gene*, **137**: 33-39.

Urlaub, H., Thiede, B., Muller, E.C., Brimarombe, R. and Wittmann-Liebold, B. (1997). *J. Biol. Chem.*, **272**: 14547-14555.

Vishwanath, R.I., et al. (1999). *Science,* **283**: 83-87.

Vlodavsky, I., Bar-Shavit, R., Ishai-Michaeli, R., Bashkin, P. and Fuks, Z. (1991). *Biochem. In Israel/TI135,* **16**: 268-271.

Vlodavsky, I., Folkman, J., Sullivan, R., Freidman, R., Ishai-Michaeli, R., Sasse, J. and Klagsbrun, M. (1987). *Proc. Natl. Acad. Sci.*, **84**: 2292-2296.

Voet, D and Voet, J.G. (1995). Biochemistry. J. Wiley and Sons, New York.

Williams, K.R. and Konigsberg, W.H. (1991). *Methods Enzymol.*, **208**: 516-539.

Willis, M.C., LeCuyer, K.A., Meisenheimer, K.M., Uhlenbeck, O.C. and Koch, T.H. (1994). *Nucleic Acids Res.*, **22**: 4947-4952.

Wong, D.L., Pavlovich, J.G. and Reich, N.O. (1998). *Nucleic Acids Res.*, **26**: 645-649.

Yayon, A., Klagsbrun, M., Esko, J.D., Leder, P. and Ornitze, M.D. (1991). *Cell*, **64**: 841-848.

Yoshida, K. (1977). *J. Am. Chem. Soc.*, **99**: 6111.

Yoshida, K. (1979). *J. Am. Chem. Soc.*, **101**: 2116.

Zhang, J., Cousens, L.S., Barr, P.J. and Sprang, S.R. (1991). *Proc. Natl. Acad. Sci.,* USA, **88**: 3446-3450.

Zhu, X., Komiya, H., Chirino, A., Faham, S., Fox, G.M., Arakawa, T., Hsu, B.T. and Rees, D.C. (1991). *Science,* **251**: 90-93.

Zucker, M. (1989). *Science,* **244**: 48-52.